



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : <b>C12N 15/00, 15/12, C07K 14/00, 14/435, 7/00</b></p>	<p><b>A1</b></p>	<p>(11) International Publication Number: <b>WO 99/50405</b></p> <p>(43) International Publication Date: <b>7 October 1999 (07.10.99)</b></p>
<p>(21) International Application Number: <b>PCT/US99/06946</b></p> <p>(22) International Filing Date: <b>30 March 1999 (30.03.99)</b></p> <p>(30) Priority Data:              60/080,110           31 March 1998 (31.03.98)       US              09/280,591           29 March 1999 (29.03.99)       US</p> <p>(71) Applicant: <b>GENETICS INSTITUTE, INC. [US/US]; 87 Cambridge Park Drive, Cambridge, MA 02140 (US).</b></p> <p>(72) Inventors: <b>JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). McCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LaVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). COLLINS-RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). EVANS, Cheryl; 18801 Bent Willow Circle, Germantown, MD 20874 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US). STEININGER, Robert, J. II; 100 Reed Street, Cambridge, MA 02140 (US).</b></p> <p>(74) Agent: <b>MANDRAGOURAS, Amy, A.; Lahive &amp; Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).</b></p>		<p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b></p> <p><b>Published</b> <i>With international search report.</i></p>
<p>(54) Title: <b>SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM</b></p>		
<p>(57) Abstract</p> <p>Novel polynucleotides and the proteins encoded thereby are disclosed.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MX	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5

## SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

10

This application is a continuation-in-part of provisional application Ser. No. 60/080,110, filed March 31, 1998, which is incorporated by reference herein.

15

### FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

20

### BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques  
25 clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader  
30 sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the  
35 polynucleotides encoding them that the present invention is directed.

### SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 5 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 216 to nucleotide 2174;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 339 to nucleotide 2174;
- 10 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cc359\_4 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;
- 15 (f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cc359\_4 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;
- 20 (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:2;
- 25 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ;
- (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and
- 30 (m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:1.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 216 to nucleotide 2174; the nucleotide sequence of SEQ ID NO:1 from nucleotide 339 to nucleotide 2174; the nucleotide sequence of the full-length protein coding sequence of clone cc359\_4 deposited under accession number ATCC 98715; or the  
5 nucleotide sequence of a mature protein coding sequence of clone cc359\_4 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715. In further preferred  
10 embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:2, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 321  
15 to amino acid 330 of SEQ ID NO:2.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

- 20 (a) a process comprising the steps of:
- (i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
- (aa) SEQ ID NO:1, but excluding the poly(A) tail at the  
25 3' end of SEQ ID NO:1; and
- (ab) the nucleotide sequence of the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;
- (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and
- 30 (iii) isolating the DNA polynucleotides detected with the probe(s);
- and
- (b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:1, but excluding the poly(A) tail at the 3' end of SEQ ID NO:1; and

(bb) the nucleotide sequence of the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:1, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:1 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:1, but excluding the poly(A) tail at the 3' end of SEQ ID NO:1. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:1 from nucleotide 216 to nucleotide 2174, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:1 from nucleotide 216 to nucleotide 2174, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:1 from nucleotide 216 to nucleotide 2174. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:1 from nucleotide 339 to nucleotide 2174, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:1 from nucleotide 339 to nucleotide 2174, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:1 from nucleotide 339 to nucleotide 2174.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) a fragment of the amino acid sequence of SEQ ID NO:2, the fragment comprising eight contiguous amino acids of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715; the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2. In further preferred  
5 embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:2, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from  
10 amino acid 321 to amino acid 330 of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- 15 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 414;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 370 to nucleotide 414;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ct547\_2 deposited under accession  
20 number ATCC 98715;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715;
- (f) a polynucleotide comprising the nucleotide sequence of a mature  
25 protein coding sequence of clone ct547\_2 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715;
- (h) a polynucleotide encoding a protein comprising the amino acid  
30 sequence of SEQ ID NO:4;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ;

5 (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:3.

10 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 414; the nucleotide sequence of SEQ ID NO:3 from nucleotide 370 to nucleotide 414; the nucleotide sequence of the full-length protein coding sequence of clone ct547\_2 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone ct547\_2 deposited  
15 under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological  
20 activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:4, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 30 to amino acid 39 of SEQ ID NO:4.

25 Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

30 (i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:3, but excluding the poly(A) tail at the 3' end of SEQ ID NO:3; and



- (ab) the nucleotide sequence of the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715;
- (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and
- 5 (iii) isolating the DNA polynucleotides detected with the probe(s);
- and
- (b) a process comprising the steps of:
- (i) preparing one or more polynucleotide primers that
- 10 hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
- (ba) SEQ ID NO:3, but excluding the poly(A) tail at the 3' end of SEQ ID NO:3; and
- (bb) the nucleotide sequence of the cDNA insert of clone
- 15 ct547\_2 deposited under accession number ATCC 98715;
- (ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;
- (iii) amplifying human DNA sequences; and
- (iv) isolating the polynucleotide products of step (b)(iii).
- 20 Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:3 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:3, but excluding the poly(A) tail at the 3' end of SEQ ID NO:3. Also preferably the polynucleotide isolated
- 25 according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 414, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 414, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:3 from nucleotide 202 to
- 30 nucleotide 414. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:3 from nucleotide 370 to nucleotide 414, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:3 from

nucleotide 370 to nucleotide 414, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:3 from nucleotide 370 to nucleotide 414.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group

5 consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
- (b) a fragment of the amino acid sequence of SEQ ID NO:4, the fragment comprising eight contiguous amino acids of SEQ ID NO:4; and

(c) the amino acid sequence encoded by the cDNA insert of clone  
10 ct547\_2 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably  
15 comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:4, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 30 to amino acid 39 of SEQ ID NO:4.

In one embodiment, the present invention provides a composition comprising an  
20 isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 206 to nucleotide 415;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID  
25 NO:5 from nucleotide 293 to nucleotide 415;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone en553\_1 deposited under accession number ATCC 98715;

(e) a polynucleotide encoding the full-length protein encoded by the  
30 cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone en553\_1 deposited under accession number ATCC 98715;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

5 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

10 (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ;

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and

15 (m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:5.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 206 to nucleotide 415; the nucleotide sequence of SEQ ID NO:5 from nucleotide 293 to nucleotide 415; the nucleotide sequence of the full-length protein coding  
20 sequence of clone en553\_1 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone en553\_1 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715. In further preferred  
25 embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:6, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having  
30 biological activity, the fragment comprising the amino acid sequence from amino acid 30 to amino acid 39 of SEQ ID NO:6.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:5, but excluding the poly(A) tail at the 3' end of SEQ ID NO:5; and

(ab) the nucleotide sequence of the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

(ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:5, but excluding the poly(A) tail at the 3' end of SEQ ID NO:5; and

(bb) the nucleotide sequence of the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:5, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:5 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:5, but excluding the poly(A) tail at the 3' end of SEQ ID NO:5. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:5 from nucleotide 206 to nucleotide 415, and extending

contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:5 from nucleotide 206 to nucleotide 415, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:5 from nucleotide 206 to nucleotide 415. Also preferably the polynucleotide isolated according to the above  
5 process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:5 from nucleotide 293 to nucleotide 415, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:5 from nucleotide 293 to nucleotide 415, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:5 from nucleotide 293 to nucleotide 415.

10 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:6;
- (b) a fragment of the amino acid sequence of SEQ ID NO:6, the  
15 fragment comprising eight contiguous amino acids of SEQ ID NO:6; and
- (c) the amino acid sequence encoded by the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6. In further preferred  
20 embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:6, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from  
25 amino acid 30 to amino acid 39 of SEQ ID NO:6.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID  
30 NO:7 from nucleotide 120 to nucleotide 1202;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone nn296\_2 deposited under accession number ATCC 98715;

- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone nn296\_2 deposited under accession number ATCC 98715;
- (e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone nn296\_2 deposited under accession number ATCC 98715;
- (f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone nn296\_2 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:8;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ;
- (k) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h); and
- (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h) and that has a length that is at least 25% of the length of SEQ ID NO:7.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 120 to nucleotide 1202; the nucleotide sequence of the full-length protein coding sequence of clone nn296\_2 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone nn296\_2 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone nn296\_2 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:8, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having

biological activity, the fragment comprising the amino acid sequence from amino acid 175 to amino acid 184 of SEQ ID NO:8.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

5 Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize  
10 in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:7, but excluding the poly(A) tail at the  
3' end of SEQ ID NO:7; and

(ab) the nucleotide sequence of the cDNA insert of clone  
nn296\_2 deposited under accession number ATCC 98715;

15 (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

20 (b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:7, but excluding the poly(A) tail at the  
25 3' end of SEQ ID NO:7; and

(bb) the nucleotide sequence of the cDNA insert of clone  
nn296\_2 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

30 (iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:7, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:7 to

a nucleotide sequence corresponding to the 3' end of SEQ ID NO:7, but excluding the poly(A) tail at the 3' end of SEQ ID NO:7. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:7 from nucleotide 120 to nucleotide 1202, and extending  
5 contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:7 from nucleotide 120 to nucleotide 1202, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:7 from nucleotide 120 to nucleotide 1202.

In other embodiments, the present invention provides a composition comprising  
10 a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:8;
- (b) a fragment of the amino acid sequence of SEQ ID NO:8, the fragment comprising eight contiguous amino acids of SEQ ID NO:8; and  
15 (c) the amino acid sequence encoded by the cDNA insert of clone nn296\_2 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8. In further preferred  
20 embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:8, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 175 to amino acid 184 of SEQ ID NO:8.

25 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID  
30 NO:9 from nucleotide 597 to nucleotide 704;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone nq27\_13 deposited under accession number ATCC 98715;



- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;
- (e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone nq27\_13 deposited under accession number ATCC 98715;
- (f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:10;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:32;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:32 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:32;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (g)-(j) above ;
- (m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j); and
- (n) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j) and that has a length that is at least 25% of the length of SEQ ID NO:9.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 597 to nucleotide 704; the nucleotide sequence of the full-length protein coding sequence of clone nq27\_13 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone nq27\_13 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein

comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:10, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 13 to amino acid 22 of SEQ ID NO:10.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:9, but excluding the poly(A) tail at the 3' end of SEQ ID NO:9; and

(ab) the nucleotide sequence of the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;

(ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:9, but excluding the poly(A) tail at the 3' end of SEQ ID NO:9; and

(bb) the nucleotide sequence of the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:9, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:9 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:9, but excluding the poly(A) tail at the 3' end of SEQ ID NO:9. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:9 from nucleotide 597 to nucleotide 704, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:9 from nucleotide 597 to nucleotide 704, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:9 from nucleotide 597 to nucleotide 704.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:10;
- (b) a fragment of the amino acid sequence of SEQ ID NO:10, the fragment comprising eight contiguous amino acids of SEQ ID NO:10; and
- (c) the amino acid sequence of SEQ ID NO:32;
- (d) a fragment of the amino acid sequence of SEQ ID NO:32, the fragment comprising eight contiguous amino acids of SEQ ID NO:32; and
- (e) the amino acid sequence encoded by the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:10, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 13 to amino acid 22 of SEQ ID NO:10.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 44 to nucleotide 475;
- 5 (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pk65\_4 deposited under accession number ATCC 98715;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;
- 10 (e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone pk65\_4 deposited under accession number ATCC 98715;
- (f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;
- 15 (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:12;
- 20 (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ;
- (k) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h); and
- 25 (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h) and that has a length that is at least 25% of the length of SEQ ID NO:11.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 44 to nucleotide 475; the nucleotide sequence of the full-length protein coding sequence of clone pk65\_4 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone pk65\_4 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert

30

of clone pk65\_4 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:12, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 67 to amino acid 76 of SEQ ID NO:12.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

- (a) a process comprising the steps of:
  - (i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
    - (aa) SEQ ID NO:11, but excluding the poly(A) tail at the 3' end of SEQ ID NO:11; and
    - (ab) the nucleotide sequence of the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;
  - (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and
  - (iii) isolating the DNA polynucleotides detected with the probe(s);
- and
- (b) a process comprising the steps of:
  - (i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
    - (ba) SEQ ID NO:11, but excluding the poly(A) tail at the 3' end of SEQ ID NO:11; and
    - (bb) the nucleotide sequence of the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

- 5 Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:11, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:11 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:11, but excluding the poly(A) tail at the 3' end of SEQ ID NO:11. Also preferably the
- 10 polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:11 from nucleotide 44 to nucleotide 475, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:11 from nucleotide 44 to nucleotide 475, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:11 from nucleotide
- 15 44 to nucleotide 475.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:12;
- 20 (b) a fragment of the amino acid sequence of SEQ ID NO:12, the fragment comprising eight contiguous amino acids of SEQ ID NO:12; and
- (c) the amino acid sequence encoded by the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such

25 protein comprises the amino acid sequence of SEQ ID NO:12. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:12, or a protein comprising a fragment of the amino acid sequence of SEQ

30 ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 67 to amino acid 76 of SEQ ID NO:12.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 285 to nucleotide 590;
- 5 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 408 to nucleotide 590;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pk855\_1 deposited under accession number ATCC 98715;
- 10 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;
- (f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone pk855\_1 deposited under accession number ATCC 98715;
- 15 (g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:14;
- 20 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ;
- 25 (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and
- (m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:13.
- 30

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 285 to nucleotide 590; the nucleotide sequence of SEQ ID NO:13 from nucleotide 408 to nucleotide 590; the nucleotide sequence of the full-length protein coding sequence of clone pk855\_1 deposited under accession number ATCC 98715; or the

nucleotide sequence of a mature protein coding sequence of clone pk855\_1 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715. In further preferred  
5 embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:14, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having  
10 biological activity, the fragment comprising the amino acid sequence from amino acid 46 to amino acid 55 of SEQ ID NO:14.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

Further embodiments of the invention provide isolated polynucleotides produced  
15 according to a process selected from the group consisting of:

(a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

20 (aa) SEQ ID NO:13, but excluding the poly(A) tail at the 3' end of SEQ ID NO:13; and

(ab) the nucleotide sequence of the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;

25 (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

30 (i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:13, but excluding the poly(A) tail at the 3' end of SEQ ID NO:13; and



- (bb) the nucleotide sequence of the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;
- (ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;
- 5 (iii) amplifying human DNA sequences; and
- (iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:13, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:13 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:13, but excluding the poly(A) tail at the 3' end of SEQ ID NO:13. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:13 from nucleotide 285 to nucleotide 590, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:13 from nucleotide 285 to nucleotide 590, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:13 from nucleotide 285 to nucleotide 590. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:13 from nucleotide 408 to nucleotide 590, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:13 from nucleotide 408 to nucleotide 590, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:13 from nucleotide 408 to nucleotide 590.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) a fragment of the amino acid sequence of SEQ ID NO:14, the fragment comprising eight contiguous amino acids of SEQ ID NO:14; and
- (c) the amino acid sequence encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment preferably

comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:14, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 46 to amino acid 55 of SEQ ID NO:14.

5        In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a)    a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b)    a polynucleotide comprising the nucleotide sequence of SEQ ID  
10    NO:15 from nucleotide 193 to nucleotide 2223;
- (c)    a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1045 to nucleotide 2223;
- (d)    a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone PL776\_6 deposited under accession  
15    number ATCC 98715;
- (e)    a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;
- (f)    a polynucleotide comprising the nucleotide sequence of a mature  
20    protein coding sequence of clone PL776\_6 deposited under accession number ATCC 98715;
- (g)    a polynucleotide encoding a mature protein encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;
- (h)    a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;
- (i)    a polynucleotide encoding a protein comprising a fragment of the  
25    amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:16;
- (j)    a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k)    a polynucleotide which encodes a species homologue of the protein  
30    of (h) or (i) above ;
- (l)    a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:15.

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 193 to nucleotide 2223; the nucleotide sequence of SEQ ID NO:15 from nucleotide 1045 to nucleotide 2223; the nucleotide sequence of the full-length protein coding sequence of clone PL776\_6 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone PL776\_6 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:16, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 333 to amino acid 342 of SEQ ID NO:16.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

- (a) a process comprising the steps of:
    - (i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
      - (aa) SEQ ID NO:15; and
      - (ab) the nucleotide sequence of the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;
    - (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and
    - (iii) isolating the DNA polynucleotides detected with the probe(s);
- and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

5

(ba) SEQ ID NO:15; and

(bb) the nucleotide sequence of the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

10

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:15, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:15 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:15. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:15 from nucleotide 193 to nucleotide 2223, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:15 from nucleotide 193 to nucleotide 2223, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:15 from nucleotide 193 to nucleotide 2223. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:15 from nucleotide 1045 to nucleotide 2223, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:15 from nucleotide 1045 to nucleotide 2223, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:15 from nucleotide 1045 to nucleotide 2223.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

30

(a) the amino acid sequence of SEQ ID NO:16;

(b) a fragment of the amino acid sequence of SEQ ID NO:16, the fragment comprising eight contiguous amino acids of SEQ ID NO:16; and

(c) the amino acid sequence encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715; the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16. In further preferred  
5 embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:16, or a protein comprising a fragment of the amino acid sequence of SEQ  
10 ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 333 to amino acid 342 of SEQ ID NO:16.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;
- 15 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 198 to nucleotide 1121;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 381 to nucleotide 1121;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pm4\_13 deposited under accession  
20 number ATCC 98715;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;
- (f) a polynucleotide comprising the nucleotide sequence of a mature  
25 protein coding sequence of clone pm4\_13 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;
- (h) a polynucleotide encoding a protein comprising the amino acid  
30 sequence of SEQ ID NO:18;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:18;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ;

5 (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i); and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i) and that has a length that is at least 25% of the length of SEQ ID NO:17.

10 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 198 to nucleotide 1121; the nucleotide sequence of SEQ ID NO:17 from nucleotide 381 to nucleotide 1121; the nucleotide sequence of the full-length protein coding sequence of clone pm4\_13 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone pm4\_13 deposited  
15 under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological  
20 activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:18, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 149 to amino acid 158 of SEQ ID NO:18.

25 Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

- (a) a process comprising the steps of:
- 30 (i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
- (aa) SEQ ID NO:17, but excluding the poly(A) tail at the 3' end of SEQ ID NO:17; and

- (ab) the nucleotide sequence of the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;
- (ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and
- 5 (iii) isolating the DNA polynucleotides detected with the probe(s);
- and
- (b) a process comprising the steps of:
- (i) preparing one or more polynucleotide primers that
- 10 hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:
- (ba) SEQ ID NO:17, but excluding the poly(A) tail at the 3' end of SEQ ID NO:17; and
- (bb) the nucleotide sequence of the cDNA insert of clone
- 15 pm4\_13 deposited under accession number ATCC 98715;
- (ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;
- (iii) amplifying human DNA sequences; and
- (iv) isolating the polynucleotide products of step (b)(iii).
- 20 Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:17, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:17 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:17, but excluding the poly(A) tail at the 3' end of SEQ ID NO:17. Also preferably the
- 25 polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:17 from nucleotide 198 to nucleotide 1121, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:17 from nucleotide 198 to nucleotide 1121, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:17 from nucleotide
- 30 198 to nucleotide 1121. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:17 from nucleotide 381 to nucleotide 1121, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:17 from

nucleotide 381 to nucleotide 1121, to a nucleotide sequence corresponding to the 3' end of said sequence of SEQ ID NO:17 from nucleotide 381 to nucleotide 1121.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
  - (b) a fragment of the amino acid sequence of SEQ ID NO:18, the fragment comprising eight contiguous amino acids of SEQ ID NO:18; and
  - (c) the amino acid sequence encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:18, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 149 to amino acid 158 of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 19 to nucleotide 1953;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pt326\_4 deposited under accession number ATCC 98715;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;
- (e) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone pt326\_4 deposited under accession number ATCC 98715;
- (f) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;



- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising eight contiguous amino acids of SEQ ID NO:20;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ;
- (k) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h); and
- (l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(h) and that has a length that is at least 25% of the length of SEQ ID NO:19.
- Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 19 to nucleotide 1953; the nucleotide sequence of the full-length protein coding sequence of clone pt326\_4 deposited under accession number ATCC 98715; or the nucleotide sequence of a mature protein coding sequence of clone pt326\_4 deposited under accession number ATCC 98715. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:20, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 317 to amino acid 326 of SEQ ID NO:20.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

Further embodiments of the invention provide isolated polynucleotides produced according to a process selected from the group consisting of:

- (a) a process comprising the steps of:

(i) preparing one or more polynucleotide probes that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(aa) SEQ ID NO:19, but excluding the poly(A) tail at the 3' end of SEQ ID NO:19; and

(ab) the nucleotide sequence of the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;

(ii) hybridizing said probe(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C; and

(iii) isolating the DNA polynucleotides detected with the probe(s);

and

(b) a process comprising the steps of:

(i) preparing one or more polynucleotide primers that hybridize in 6X SSC at 65 degrees C to a nucleotide sequence selected from the group consisting of:

(ba) SEQ ID NO:19, but excluding the poly(A) tail at the 3' end of SEQ ID NO:19; and

(bb) the nucleotide sequence of the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;

(ii) hybridizing said primer(s) to human genomic DNA in conditions at least as stringent as 4X SSC at 50 degrees C;

(iii) amplifying human DNA sequences; and

(iv) isolating the polynucleotide products of step (b)(iii).

Preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:19, and extending contiguously from a nucleotide sequence corresponding to the 5' end of SEQ ID NO:19 to a nucleotide sequence corresponding to the 3' end of SEQ ID NO:19, but excluding the poly(A) tail at the 3' end of SEQ ID NO:19. Also preferably the polynucleotide isolated according to the above process comprises a nucleotide sequence corresponding to the cDNA sequence of SEQ ID NO:19 from nucleotide 19 to nucleotide 1953, and extending contiguously from a nucleotide sequence corresponding to the 5' end of said sequence of SEQ ID NO:19 from nucleotide 19 to nucleotide 1953, to a nucleotide

sequence corresponding to the 3' end of said sequence of SEQ ID NO:19 from nucleotide 19 to nucleotide 1953.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:20;
- (b) a fragment of the amino acid sequence of SEQ ID NO:20, the fragment comprising eight contiguous amino acids of SEQ ID NO:20; and
- (c) the amino acid sequence encoded by the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) contiguous amino acids of SEQ ID NO:20, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 317 to amino acid 326 of SEQ ID NO:20.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

- (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and
- (b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

5

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

10

#### DETAILED DESCRIPTION

##### ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

##### Clone "cc359\_4"

A polynucleotide of the present invention has been identified as clone "cc359\_4". cc359\_4 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cc359\_4 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "cc359\_4 protein").

The nucleotide sequence of cc359\_4 as presently determined is reported in SEQ ID NO:1, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cc359\_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 29 to 41 of SEQ ID NO:2 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 42. Due to the hydrophobic nature of the predicted leader/signal sequence, it is likely to act as a transmembrane domain should the predicted leader/signal sequence not be separated from the remainder of the cc359\_4 protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cc359\_4 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for cc359\_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cc359\_4 demonstrated at least some similarity with sequences identified as H23117 (ym51g12.s1 Homo sapiens cDNA clone 51970 3'). Based upon sequence similarity, cc359\_4 proteins and each similar protein or peptide may share at least some activity. The cc359\_4 protein sequence contains four chaoptin domains (at amino acids 94-115, 118-139, 142-163, and 261-282 of SEQ ID NO:2). Chaoptin is a cell-surface glycoprotein required for *Drosophila* photoreceptor cell morphogenesis. Chaoptin is largely composed of 41 potentially amphipathic repeats. Chaoptin repeats that have been reported in both yeast and human proteins are remarkably similar, suggesting their general importance as a structural and/or functional motif. Additionally, the cc359\_4 protein was found to contain a leucine zipper motif. The TopPredII computer program predicts five potential transmembrane domains within the cc359\_4 protein sequence, centered around amino acids 20, 410, 490, 530, and 590 of SEQ ID NO:2, respectively.

#### Clone "ct547\_2"

A polynucleotide of the present invention has been identified as clone "ct547\_2". ct547\_2 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ct547\_2 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "ct547\_2 protein").

The nucleotide sequence of ct547\_2 as presently determined is reported in SEQ ID NO:3, and includes a poly(A) tail. What applicants presently believe to be the proper  
5 reading frame and the predicted amino acid sequence of the ct547\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 44 to 56 of SEQ ID NO:4 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 57. Due to the hydrophobic nature of the predicted leader/signal sequence, it is likely to act as a transmembrane domain  
10 should the predicted leader/signal sequence not be separated from the remainder of the ct547\_2 protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ct547\_2 should be approximately 1600 bp.

The nucleotide sequence disclosed herein for ct547\_2 was searched against the  
15 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ct547\_2 demonstrated at least some similarity with sequences identified as AA428546 (zw47d03.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 773189 5'), R20032 (yg31g04.r1 Homo sapiens cDNA clone 34069 5'), R27350 (yh53c12.r1 Homo sapiens cDNA clone 133462 5'), and R66971 (yi29e12.r1 Homo sapiens  
20 cDNA clone 140686 5'). Based upon sequence similarity, ct547\_2 proteins and each similar protein or peptide may share at least some activity.

#### Clone "en553\_1"

A polynucleotide of the present invention has been identified as clone "en553\_1".  
25 en553\_1 was isolated from a human fetal brain cDNA library and was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. en553\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "en553\_1 protein").

The nucleotide sequence of en553\_1 as presently determined is reported in SEQ  
30 ID NO:5, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the en553\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 17 to 29 of SEQ ID NO:6 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30. Due to the hydrophobic nature

of the predicted leader/signal sequence, it is likely to act as a transmembrane domain should the predicted leader/signal sequence not be separated from the remainder of the en553\_1 protein. Another potential en553\_1 reading frame and predicted amino acid sequence is encoded by basepairs 1334 to 1597 of SEQ ID NO:5 and is reported in SEQ ID NO:31

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone en553\_1 should be approximately 2675 bp.

The nucleotide sequence disclosed herein for en553\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. en553\_1 demonstrated at least some similarity with sequences identified as No hits were found in the databases. The nucleotide sequence of en553\_1 indicates that it may contain one or more copies of the following repetitive elements: Alu, MIR.

#### Clone "nn296\_2"

A polynucleotide of the present invention has been identified as clone "nn296\_2". nn296\_2 was isolated from a human fetal kidney (293 cell line) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. nn296\_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "nn296\_2 protein").

The nucleotide sequence of nn296\_2 as presently determined is reported in SEQ ID NO:7, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the nn296\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone nn296\_2 should be approximately 4500 bp.

The nucleotide sequence disclosed herein for nn296\_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. nn296\_2 demonstrated at least some similarity with sequences identified as AA054538 (zk83e02.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 489434.3), AA175407 (ms80c06.r1 Soares mouse 3NbMS Mus musculus cDNA clone 617866 5' similar to WP R08C7.2 CE07425), AA510586 (vg33c05.r1 Soares mouse

mammary gland NbMMG Mus musculus cDNA clone 863144 5' similar to WP:R08C7.2 CE07425), C18077 (Human placenta cDNA 5'-end GEN-557B03), T24416 (Human gene signature HUMGS06449), and W74238 (zd02f09.r1 Pancreatic Islet Homo sapiens cDNA clone 339497 5'). The predicted amino acid sequence disclosed herein for nn296\_2 was  
5 searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted nn296\_2 protein demonstrated at least some similarity to sequences identified as U61953 (cosmid R08C7 [Caenorhabditis elegans]). Based upon sequence similarity, nn296\_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential  
10 transmembrane domains within the nn296\_2 protein sequence, one centered around amino acid 50 and another around amino acid 340 of SEQ ID NO:8. The nucleotide sequence of nn296\_2 indicates that it may contain at least one repetitive element.

#### Clone "nq27\_13"

15 A polynucleotide of the present invention has been identified as clone "nq27\_13". nq27\_13 was isolated from a human adult blood (erythroleukemia TF-1) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. nq27\_13  
20 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "nq27\_13 protein").

The nucleotide sequence of nq27\_13 as presently determined is reported in SEQ ID NO:9, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the nq27\_13 protein  
25 corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Another potential nq27\_13 reading frame and predicted amino acid sequence is encoded by basepairs 371 to 616 of SEQ ID NO:9 and is reported in SEQ ID NO:32.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone nq27\_13 should be approximately 900 bp.

30 The nucleotide sequence disclosed herein for nq27\_13 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. nq27\_13 demonstrated at least some similarity with sequences identified as AA015599 (ze20h10.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 359587 3'), AA465112 (aa32c05.r1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone



IMAGE:814952 5'), AA628397 (af26a12.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 1032766 3'), AA628438 (af26e12.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 1032814 3'), D79280 (Human aorta cDNA 5'-end GEN-213H05), and H81836 (ys68g11.r1 Homo sapiens cDNA clone 220004 5'). Based upon sequence  
 5 similarity, nq27\_13 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain at the N-terminus of the nq27\_13 protein sequence, centered around amino acid 13 of SEQ ID NO:10.

nq27\_13 protein was expressed in a COS cell expression system, and an expressed  
 10 protein band of approximately 11 kDa was detected in membrane fractions using SDS polyacrylamide gel electrophoresis.

#### Clone "pk65\_4"

A polynucleotide of the present invention has been identified as clone "pk65\_4".  
 15 pk65\_4 was isolated from a human fetal kidney (293 cell line) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. pk65\_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to  
 20 herein as "pk65\_4 protein").

The nucleotide sequence of pk65\_4 as presently determined is reported in SEQ ID NO:11, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the pk65\_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12.

25 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone pk65\_4 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for pk65\_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. pk65\_4 demonstrated at least some similarity with sequences  
 30 identified as AA309041 (EST179822 Colon carcinoma (Caco-2) cell line I Homo sapiens cDNA 5' end), AA316187 (EST187903 HCC cell line (matatasis to liver in mouse) II Homo sapiens cDNA 5' end), AA488843 (aa55a10.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:824826 3'), AF022811 (Mus musculus cornichon mRNA, complete cds), D16980 (Human HepG2 partial cDNA, clone hmd2a04m5), N44081 (yy32a02.r1 Homo

sapiens cDNA clone 272906 5' similar to PIR A56724 A56724 cni protein - fruit fly). The predicted amino acid sequence disclosed herein for pk65\_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted pk65\_4 protein demonstrated at least some similarity to sequences  
5 identified as AB006191 (cornichon-like protein [Mus musculus]), AF022811 (cornichon [Mus musculus]), and U28069 (cni gene product [Drosophila melanogaster]). Based upon sequence similarity, pk65\_4 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three potential transmembrane domains within the pk65\_4 protein sequence, centered around amino  
10 acids 16, 67, and 133 of SEQ ID NO:12, respectively.

pk65\_4 protein was expressed in a COS cell expression system, and an expressed protein band of approximately 15 kDa was detected in membrane fractions using SDS polyacrylamide gel electrophoresis.

15 Clone "pk855\_1"

A polynucleotide of the present invention has been identified as clone "pk855\_1". pk855\_1 was isolated from a human fetal kidney (293 cell line) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis  
20 of computer analysis of the amino acid sequence of the encoded protein. pk855\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "pk855\_1 protein").

The nucleotide sequence of pk855\_1 as presently determined is reported in SEQ ID NO:13, and includes a poly(A) tail. What applicants presently believe to be the proper  
25 reading frame and the predicted amino acid sequence of the pk855\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Amino acids 29 to 41 of SEQ ID NO:14 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 42. Due to the hydrophobic nature of the predicted leader/signal sequence, it is likely to act as a transmembrane domain  
30 should the predicted leader/signal sequence not be separated from the remainder of the pk855\_1 protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone pk855\_1 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for pk855\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. pk855\_1 demonstrated at least some similarity with sequences identified as AA236320 (zr53h03.r1 Soares NhHMPu S1 Homo sapiens cDNA clone  
5 667157 5'), AA491296 (aa53d03.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:824645 3'), and W76249 (zd66e06.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 345634 5'). Based upon sequence similarity, pk855\_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts an additional potential transmembrane domain within the pk855\_1  
10 protein sequence centered around amino acid 78 of SEQ ID NO:14. The nucleotide sequence of pk855\_1 indicates that it may contain some repeat elements (including some short stretches of poly(T)).

#### Clone "PL776\_6"

15 A polynucleotide of the present invention has been identified as clone "PL776\_6". PL776\_6 was isolated from a human fetal kidney (293 cell line) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. PL776\_6 is a full-  
20 length clone, including the entire coding sequence of a secreted protein (also referred to herein as "PL776\_6 protein").

The nucleotide sequence of PL776\_6 as presently determined is reported in SEQ ID NO:15, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the PL776\_6 protein  
25 corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Amino acids 272 to 284 of SEQ ID NO:16 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 285. Due to the hydrophobic nature of the predicted leader/signal sequence, it is likely to act as a transmembrane domain should the predicted leader/signal sequence not be separated  
30 from the remainder of the PL776\_6 protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone PL776\_6 should be approximately 3600 bp.

The nucleotide sequence disclosed herein for PL776\_6 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and

FASTA search protocols. PL776\_6 demonstrated at least some similarity with sequences identified as AA009770 (zi04c06.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 429802 3'), AA056343 (zl66f04.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 509599 3'), AA076433 (zm19g07.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526140 3'), AA076517 (zm19g07.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526140 5'), M93661 (Rat notch 2 mRNA), and R36803 (CBS-216 Homo sapiens cDNA clone CBS-216 5' end). Based upon sequence similarity, PL776\_6 proteins and each similar protein or peptide may share at least some activity.

#### 10 Clone "pm4\_13"

A polynucleotide of the present invention has been identified as clone "pm4\_13". pm4\_13 was isolated from a human fetal kidney (293 cell line) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. pm4\_13 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "pm4\_13 protein").

The nucleotide sequence of pm4\_13 as presently determined is reported in SEQ ID NO:17, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the pm4\_13 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 49 to 61 of SEQ ID NO:18 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 62. Due to the hydrophobic nature of the predicted leader/signal sequence, it is likely to act as a transmembrane domain should the predicted leader/signal sequence not be separated from the remainder of the pm4\_13 protein.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone pm4\_13 should be approximately 3100 bp.

The nucleotide sequence disclosed herein for pm4\_13 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. pm4\_13 demonstrated at least some similarity with sequences identified as H98853 (yx15d10.s1 Homo sapiens cDNA clone 261811 3'), R43710 (yg19f12.s1 Homo sapiens cDNA clone 32692 3'), and T22962 (Human gene signature HUMGS04687). The predicted amino acid sequence disclosed herein for pm4\_13 was

searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted pm4\_13 protein demonstrated at least some similarity to sequences identified as U20865 (y1r246wp [Saccharomyces cerevisiae]), Z29121 (ZK757.1 [Caenorhabditis elegans]), Z46259 (N0325 gene product [Saccharomyces cerevisiae]), Z68507 (m18.8 [Caenorhabditis elegans]), and Z71602 (orf ynl326c [Saccharomyces cerevisiae]). Based upon sequence similarity, pm4\_13 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three additional potential transmembrane domains within the pm4\_13 protein sequence, centered around amino acids 86, 184, and 225 of SEQ ID NO:18, respectively.

#### Clone "pt326\_4"

A polynucleotide of the present invention has been identified as clone "pt326\_4". pt326\_4 was isolated from a human adult blood (lymphoblastic leukemia MOLT-4) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. pt326\_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "pt326\_4 protein").

The nucleotide sequence of pt326\_4 as presently determined is reported in SEQ ID NO:19, and includes a poly(A) tail. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the pt326\_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone pt326\_4 should be approximately 3100 bp.

The nucleotide sequence disclosed herein for pt326\_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. pt326\_4 demonstrated at least some similarity with sequences identified as AA361324 (EST70660 T-cell lymphoma Homo sapiens cDNA 5' end), H94651 (yv19c12.s1 Homo sapiens cDNA clone 243190 3'), T22309 (Human gene signature HUMGS03882), and U17907 (Human chromosome 17q21 mRNA clone B169). The predicted amino acid sequence disclosed herein for pt326\_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted pt326\_4 protein demonstrated at least some similarity to sequences identified as D89646 (GTBP-ALT [Homo sapiens]) and U28946 (G/T mismatch binding

protein (GTBP) [Homo sapiens]). Based upon sequence similarity, pt326\_4 proteins and each similar protein or peptide may share at least some activity.

#### Deposit of Clones

5 Clones cc359\_4, ct547\_2, en553\_1, nn296\_2, nq27\_13, pk65\_4, pk855\_1, PL776\_6, pm4\_13, and pt326\_4 were deposited on March 31, 1998 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original deposit under the Budapest Treaty and were given the accession number ATCC 98715, from which each clone comprising a particular polynucleotide is obtainable. All  
10 restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b), and the term of the deposit will comply with 37 C.F.R. § 1.806.

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited  
15 by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Figures 1A and 1B, respectively. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector  
20 was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce  
25 the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

30 An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of an oligonucleotide probe that was used to isolate or to sequence each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

<u>Clone</u>	<u>Probe Sequence</u>
cc359_4	SEQ ID NO:21
ct547_2	SEQ ID NO:22
en553_1	SEQ ID NO:23
5 nn296_2	SEQ ID NO:24
nq27_13	SEQ ID NO:25
pk65_4	SEQ ID NO:26
pk855_1	SEQ ID NO:27
PL776_6	SEQ ID NO:28
10 pm4_13	SEQ ID NO:29
pt326_4	SEQ ID NO:30

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a nucleotide (such as, for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
- (b) It should be designed to have a  $T_m$  of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with  $\gamma$ -<sup>32</sup>P ATP (specific activity 6000 Ci/mmmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100  $\mu$ l of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100  $\mu$ g/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in

fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other  
5 known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with  
10 NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed  
15 by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

20 The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the  
25 protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, *Bio/Technology* 10, 773-778 (1992) and in R.S. McDowell, *et al.*, *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding  
30 sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.



The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length  
5 polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that  
10 are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can  
15 be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of  
20 the organism from which the gene was isolated.

The chromosomal location corresponding to the polynucleotide sequences disclosed herein may also be determined, for example by hybridizing appropriately labeled polynucleotides of the present invention to chromosomes *in situ*. It may also be possible to determine the corresponding chromosomal location for a disclosed  
25 polynucleotide by identifying significantly similar nucleotide sequences in public databases, such as expressed sequence tags (ESTs), that have already been mapped to particular chromosomal locations. For at least some of the polynucleotide sequences disclosed herein, public database sequences having at least some similarity to the polynucleotide of the present invention have been listed by database accession number.  
30 Searches using the GenBank accession numbers of these public database sequences can then be performed at an Internet site provided by the National Center for Biotechnology Information having the address <http://www.ncbi.nlm.nih.gov/UniGene/>, in order to identify "UniGene clusters" of overlapping sequences. Many of the "UniGene clusters" so identified will already have been mapped to particular chromosomal sites.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* 15(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* 62(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* 58: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* 14(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms, part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with

known techniques for determination of such domains from sequence information. For example, the TopPredIII computer program can be used to predict the location of transmembrane domains in an amino acid sequence, domains which are described by the location of the center of the transmembrane domain, with at least ten transmembrane  
5 amino acids on each side of the reported central residue(s).

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95%  
10 identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that  
15 shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

In particular, sequence identity may be determined using WU-BLAST (Washington University BLAST) version 2.0 software, which builds upon WU-BLAST version 1.4, which in turn is based on the public domain NCBI-BLAST version 1.4  
20 (Altschul and Gish, 1996, Local alignment statistics, Doolittle *ed.*, *Methods in Enzymology* 266: 460-480; Altschul *et al.*, 1990, Basic local alignment search tool, *Journal of Molecular Biology* 215: 403-410; Gish and States, 1993, Identification of protein coding regions by database similarity search, *Nature Genetics* 3: 266-272; Karlin and Altschul, 1993, Applications and statistics for multiple high-scoring segments in molecular  
25 sequences, *Proc. Natl. Acad. Sci. USA* 90: 5873-5877; all of which are incorporated by reference herein). WU-BLAST version 2.0 executable programs for several UNIX platforms can be downloaded from <ftp://blast.wustl.edu/blast/executables>. The complete suite of search programs (BLASTP, BLASTN, BLASTX, TBLASTN, and TBLASTX) is provided at that site, in addition to several support programs. WU-BLAST 2.0 is  
30 copyrighted and may not be sold or redistributed in any form or manner without the express written consent of the author; but the posted executables may otherwise be freely used for commercial, nonprofit, or academic purposes. In all search programs in the suite

-- BLASTP, BLASTN, BLASTX, TBLASTN and TBLASTX -- the gapped alignment routines are integral to the database search itself, and thus yield much better sensitivity and selectivity while producing the more easily interpreted output. Gapping can optionally be turned off in all of these programs, if desired. The default penalty (Q) for a gap of length one is Q=9 for proteins and BLASTP, and Q=10 for BLASTN, but may be changed to any integer value including zero, one through eight, nine, ten, eleven, twelve through twenty, twenty-one through fifty, fifty-one through one hundred, etc. The default per-residue penalty for extending a gap (R) is R=2 for proteins and BLASTP, and R=10 for BLASTN, but may be changed to any integer value including zero, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve through twenty, twenty-one through fifty, fifty-one through one hundred, etc. Any combination of values for Q and R can be used in order to align sequences so as to maximize overlap and identity while minimizing sequence gaps. The default amino acid comparison matrix is BLOSUM62, but other amino acid comparison matrices such as PAM can be utilized.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*, *Hylobates concolor*, *Macaca mulatta*, *Papio papio*, *Papio hamadryas*, *Cercopithecus aethiops*, *Cebus capucinus*, *Aotus trivirgatus*, *Sanguinus oedipus*, *Microcebus murinus*, *Mus musculus*, *Rattus norvegicus*,

*Cricetulus griseus*, *Felis catus*, *Mustela vison*, *Canis familiaris*, *Oryctolagus cuniculus*, *Bos taurus*, *Ovis aries*, *Sus scrofa*, and *Equus caballus*, for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species

5 (O'Brien and Seuáñez, 1988, *Ann. Rev. Genet.* 22: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* 25: 682-690; Lyons *et al.*, 1997, *Nature Genetics* 15: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* 13(10): 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, where sequence identity is determined by

10 comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

15

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

20

The present invention also includes polynucleotides that hybridize under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

25

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) <sup>‡</sup>	Hybridization Temperature and Buffer <sup>†</sup>	Wash Temperature and Buffer <sup>†</sup>
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T <sub>B</sub> <sup>*</sup> ; 1xSSC	T <sub>B</sub> <sup>*</sup> ; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T <sub>D</sub> <sup>*</sup> ; 1xSSC	T <sub>D</sub> <sup>*</sup> ; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T <sub>F</sub> <sup>*</sup> ; 1xSSC	T <sub>F</sub> <sup>*</sup> ; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T <sub>H</sub> <sup>*</sup> ; 4xSSC	T <sub>H</sub> <sup>*</sup> ; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T <sub>J</sub> <sup>*</sup> ; 4xSSC	T <sub>J</sub> <sup>*</sup> ; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T <sub>L</sub> <sup>*</sup> ; 2xSSC	T <sub>L</sub> <sup>*</sup> ; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T <sub>N</sub> <sup>*</sup> ; 6xSSC	T <sub>N</sub> <sup>*</sup> ; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T <sub>P</sub> <sup>*</sup> ; 6xSSC	T <sub>P</sub> <sup>*</sup> ; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T <sub>R</sub> <sup>*</sup> ; 4xSSC	T <sub>R</sub> <sup>*</sup> ; 4xSSC

<sup>‡</sup>: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

<sup>†</sup>: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

<sup>\*</sup>T<sub>B</sub> - T<sub>R</sub>: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T<sub>m</sub>) of the hybrid, where T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length, T<sub>m</sub>(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T<sub>m</sub>(°C) = 81.5 + 16.6(log<sub>10</sub>[Na<sup>+</sup>]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds.,  
5 John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or  
10 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an  
15 expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably  
20 linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the  
25 protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

30 Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLabs (Beverly, MA), Pharmacia (Piscataway, NJ) and Invitrogen Corporation (Carlsbad, CA), respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from the Eastman Kodak Company (New Haven, CT).



Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art

given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

### USES AND BIOLOGICAL ACTIVITY

5       The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies  
10       or vectors suitable for introduction of DNA).

#### Research Uses and Utilities

      The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express  
15       recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare  
20       with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for  
25       examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those  
30       described in Gyuris *et al.*, 1993, *Cell* 75: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine  
5 levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the  
10 protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent  
15 grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to  
20 Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

#### Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein  
25 or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention  
30 can be added to the medium in or on which the microorganism is cultured.

#### Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may

induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is  
 5 evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

10 Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986;  
 15 Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell  
 20 stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon  $\gamma$ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic  
 25 cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature* 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983;  
 30 Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991;

Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

#### Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, *Leishmania* spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease.

Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to regulate immune responses in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term

tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

5       The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as  
10 described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

15       Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms.  
20 Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from  
25 the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and  
30 murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune

response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I  $\alpha$  chain protein and  $\beta_2$



microglobulin protein or an MHC class II  $\alpha$  chain protein and an MHC class II  $\beta$  chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated  
 5 immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated  
 10 immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without  
 15 limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, *Immunologic studies in Humans*); Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al.,  
 20 *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bowman et al., *J. Virology* 61:1992-1998; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnoli et al.,  
 25 *Cellular Immunology* 133:327-341, 1991; Brown et al., *J. Immunol.* 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J. Immunol.* 144:3028-3033, 1990; and Assays for B cell function: *In vitro*  
 30 antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek,

D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

- 5 Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995;
- 10 Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

- Assays for lymphocyte survival/apoptosis (which will identify, among others,
- 15 proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993;
- 20 Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

- Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

25

#### Hematopoiesis Regulating Activity

- A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell
- 30 lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid

cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. *Cellular Biology* 15:141-151, 1995; Keller et al., *Molecular and Cellular Biology* 13:473-486, 1993; McClanahan et al., *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., *Proc. Natl. Acad. Sci. USA* 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., *Experimental Hematology* 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and

Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

5

#### Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns,  
10 incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as  
15 well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal  
20 disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue  
25 destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in  
30 circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and

in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation

of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium ).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

#### Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin  $\alpha$  family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- $\beta$  group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

#### Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al.

APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

#### Hemostatic and Thrombolytic Activity

5 A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting  
10 formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

15 Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### Receptor/Ligand Activity

20 A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands,  
25 receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant  
30 receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:



Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160, 1989; Stoltenberg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

#### 10      Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

25

#### Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved

extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the first cadherin domain provide the basis for homophilic adhesion; modification of this  
5 recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells  
10 become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas  
15 to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed  
20 in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the  
25 inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block  
30 the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

#### Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via antibody-dependent cell-mediated cytotoxicity (ADCC)). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

#### Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s);

effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic  
5 lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another  
10 material or entity which is cross-reactive with such protein.

#### ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a  
15 pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the  
20 carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other  
25 agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor,  
30 thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical

compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in  
5 combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If  
10 administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical  
15 composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is  
20 administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention.  
25 When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid  
30 form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present

invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should  
5 contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

10 The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician  
15 will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01  $\mu$ g to about 100  
20 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1  $\mu$ g to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is  
25 contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain  
30 polyclonal and monoclonal antibodies which specifically react with the protein. As used herein, the term "antibody" includes without limitation a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a single-chain antibody, a CDR-grafted antibody, a humanized antibody, or fragments thereof which bind to the indicated protein.

Such term also includes any other species derived from an antibody or antibody sequence which is capable of binding the indicated protein.

Antibodies to a particular protein can be produced by methods well known to those skilled in the art. For example, monoclonal antibodies can be produced by generation of antibody-producing hybridomas in accordance with known methods (see for example, 5 Goding, 1983, *Monoclonal antibodies: principles and practice*, Academic Press Inc., New York; and Yokoyama, 1992, "Production of Monoclonal Antibodies" in *Current Protocols in Immunology*, Unit 2.5, Greene Publishing Assoc. and John Wiley & Sons). Polyclonal sera and antibodies can be produced by inoculation of a mammalian subject with the 10 relevant protein or fragments thereof in accordance with known methods. Fragments of antibodies, receptors, or other reactive peptides can be produced from the corresponding antibodies by cleavage of and collection of the desired fragments in accordance with known methods (see for example, Goding, *supra*; and Andrew et al., 1992, "Fragmentation of Immunoglobulins" in *Current Protocols in Immunology*, Unit 2.8, Greene Publishing 15 Assoc. and John Wiley & Sons). Chimeric antibodies and single chain antibodies can also be produced in accordance with known recombinant methods (see for example, 5,169,939, 5,194,594, and 5,576,184). Humanized antibodies can also be made from corresponding murine antibodies in accordance with well known methods (see for example, U.S. Patent Nos. 5,530,101, 5,585,089, and 5,693,762). Additionally, human antibodies may be 20 produced in non-human animals such as mice that have been genetically altered to express human antibody molecules (see for example Fishwild *et al.*, 1996, *Nature Biotechnology* 14: 845-851; Mendez *et al.*, 1997, *Nature Genetics* 15: 146-156 (erratum *Nature Genetics* 16: 410); and U.S. Patents 5,877,397 and 5,625,126). Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide 25 immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987).

Monoclonal antibodies binding to the protein of the invention may be useful 30 diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where



abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

5 For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably  
10 be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the  
15 methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical  
20 applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium  
25 sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other  
30 ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions  
5 from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of  
10 carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to  
15 provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in  
20 question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to  
25 humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of  
30 a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect

the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

5 Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

10 Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 216 to nucleotide 2174;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 339 to nucleotide 2174;
  - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cc359\_4 deposited under accession number ATCC 98715;
  - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;
  - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cc359\_4 deposited under accession number ATCC 98715;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2, the fragment comprising eight contiguous amino acids of SEQ ID NO:2;
  - (j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i); and
  - (k) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i), and that has a length that is at least 25% of the length of SEQ ID NO:1.
2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.
4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:
  - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
  - (b) purifying said protein from the culture.
6. A protein produced according to the process of claim 5.
7. An isolated polynucleotide encoding the protein of claim 6.
8. The polynucleotide of claim 7, wherein the polynucleotide comprises the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715.
9. A protein comprising an amino acid sequence selected from the group consisting of:
  - (a) the amino acid sequence of SEQ ID NO:2;
  - (b) a fragment of the amino acid sequence of SEQ ID NO:2, the fragment comprising eight contiguous amino acids of SEQ ID NO:2; and
  - (c) the amino acid sequence encoded by the cDNA insert of clone cc359\_4 deposited under accession number ATCC 98715;the protein being substantially free from other mammalian proteins.
10. The protein of claim 9, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
11. A composition comprising the protein of claim 9 and a pharmaceutically acceptable carrier.
12. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 414;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 370 to nucleotide 414;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ct547\_2 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ct547\_2 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4, the fragment comprising eight contiguous amino acids of SEQ ID NO:4;
- (j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i); and
- (k) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i), and that has a length that is at least 25% of the length of SEQ ID NO:3.

13. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
- (b) a fragment of the amino acid sequence of SEQ ID NO:4, the fragment comprising eight contiguous amino acids of SEQ ID NO:4; and

(c) the amino acid sequence encoded by the cDNA insert of clone ct547\_2 deposited under accession number ATCC 98715; the protein being substantially free from other mammalian proteins.

14. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 206 to nucleotide 415;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 293 to nucleotide 415;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone en553\_1 deposited under accession number ATCC 98715;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone en553\_1 deposited under accession number ATCC 98715;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6, the fragment comprising eight contiguous amino acids of SEQ ID NO:6;

(j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i); and

(k) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i), and that has a length that is at least 25% of the length of SEQ ID NO:5.

15. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:6;
- (b) a fragment of the amino acid sequence of SEQ ID NO:6, the fragment comprising eight contiguous amino acids of SEQ ID NO:6; and
- (c) the amino acid sequence encoded by the cDNA insert of clone en553\_1 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

16. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 120 to nucleotide 1202;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone nn296\_2 deposited under accession number ATCC 98715;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone nn296\_2 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
- (f) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8, the fragment comprising eight contiguous amino acids of SEQ ID NO:8;
- (g) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(f); and
- (h) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(f), and that has a length that is at least 25% of the length of SEQ ID NO:7.

17. A protein comprising an amino acid sequence selected from the group consisting of:



- (a) the amino acid sequence of SEQ ID NO:8;
  - (b) a fragment of the amino acid sequence of SEQ ID NO:8, the fragment comprising eight contiguous amino acids of SEQ ID NO:8; and
  - (c) the amino acid sequence encoded by the cDNA insert of clone nn296\_2 deposited under accession number ATCC 98715;
- the protein being substantially free from other mammalian proteins.

18. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 597 to nucleotide 704;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone nq27\_13 deposited under accession number ATCC 98715;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;
- (f) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10, the fragment comprising eight contiguous amino acids of SEQ ID NO:10;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:32;
- (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:32, the fragment comprising eight contiguous amino acids of SEQ ID NO:32;
- (i) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(h); and
- (j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(h), and that has a length that is at least 25% of the length of SEQ ID NO:9.

19. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:10;
- (b) a fragment of the amino acid sequence of SEQ ID NO:10, the fragment comprising eight contiguous amino acids of SEQ ID NO:10; and
- (c) the amino acid sequence of SEQ ID NO:32;
- (d) a fragment of the amino acid sequence of SEQ ID NO:32, the fragment comprising eight contiguous amino acids of SEQ ID NO:32; and
- (e) the amino acid sequence encoded by the cDNA insert of clone nq27\_13 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

20. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 44 to nucleotide 475;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pk65\_4 deposited under accession number ATCC 98715;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;
- (f) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12, the fragment comprising eight contiguous amino acids of SEQ ID NO:12;
- (g) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(f); and
- (h) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(f), and that has a length that is at least 25% of the length of SEQ ID NO:11.

21. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:12;
- (b) a fragment of the amino acid sequence of SEQ ID NO:12, the fragment comprising eight contiguous amino acids of SEQ ID NO:12; and
- (c) the amino acid sequence encoded by the cDNA insert of clone pk65\_4 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

22. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 285 to nucleotide 590;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 408 to nucleotide 590;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pk855\_1 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone pk855\_1 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14, the fragment comprising eight contiguous amino acids of SEQ ID NO:14;
- (j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i); and

(k) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i), and that has a length that is at least 25% of the length of SEQ ID NO:13.

23. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;
- (b) a fragment of the amino acid sequence of SEQ ID NO:14, the fragment comprising eight contiguous amino acids of SEQ ID NO:14; and
- (c) the amino acid sequence encoded by the cDNA insert of clone pk855\_1 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

24. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 193 to nucleotide 2223;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1045 to nucleotide 2223;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone PL776\_6 deposited under accession number ATCC 98715;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone PL776\_6 deposited under accession number ATCC 98715;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16, the fragment comprising eight contiguous amino acids of SEQ ID NO:16;

(j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i); and

(k) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i), and that has a length that is at least 25% of the length of SEQ ID NO:15.

25. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) a fragment of the amino acid sequence of SEQ ID NO:16, the fragment comprising eight contiguous amino acids of SEQ ID NO:16; and

(c) the amino acid sequence encoded by the cDNA insert of clone PL776\_6 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

26. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 198 to nucleotide 1121;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 381 to nucleotide 1121;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pm4\_13 deposited under accession number ATCC 98715;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone pm4\_13 deposited under accession number ATCC 98715;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18, the fragment comprising eight contiguous amino acids of SEQ ID NO:18;

(j) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i); and

(k) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(i), and that has a length that is at least 25% of the length of SEQ ID NO:17.

27. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) a fragment of the amino acid sequence of SEQ ID NO:18, the fragment comprising eight contiguous amino acids of SEQ ID NO:18; and

(c) the amino acid sequence encoded by the cDNA insert of clone pm4\_13 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

28. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 19 to nucleotide 1953;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone pt326\_4 deposited under accession number ATCC 98715;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;

(e) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(f) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20, the fragment comprising eight contiguous amino acids of SEQ ID NO:20;

(g) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 65 degrees C, or 4X SSC at 42 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(f); and

(h) a polynucleotide that hybridizes under conditions at least as stringent as 4X SSC at 50 degrees C, or 6X SSC at 40 degrees C with 50% formamide, to any one of the polynucleotides specified in (a)-(f), and that has a length that is at least 25% of the length of SEQ ID NO:19.

29. A protein comprising an amino acid sequence selected from the group consisting of:

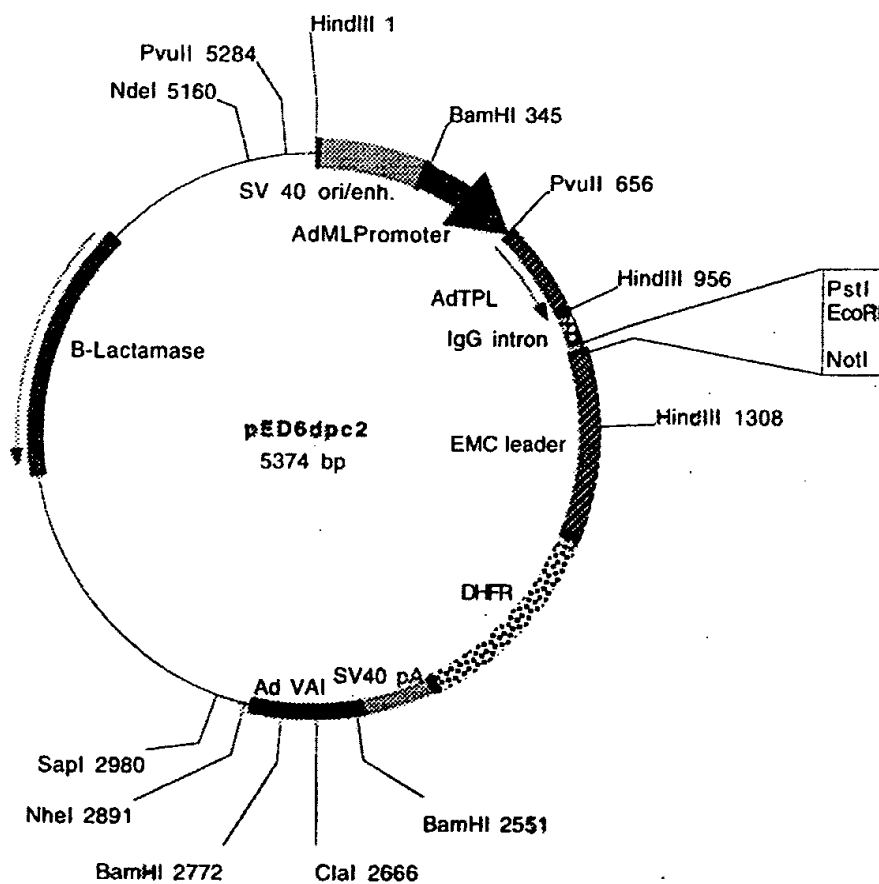
(a) the amino acid sequence of SEQ ID NO:20;

(b) a fragment of the amino acid sequence of SEQ ID NO:20, the fragment comprising eight contiguous amino acids of SEQ ID NO:20; and

(c) the amino acid sequence encoded by the cDNA insert of clone pt326\_4 deposited under accession number ATCC 98715;

the protein being substantially free from other mammalian proteins.

FIGURE 1A



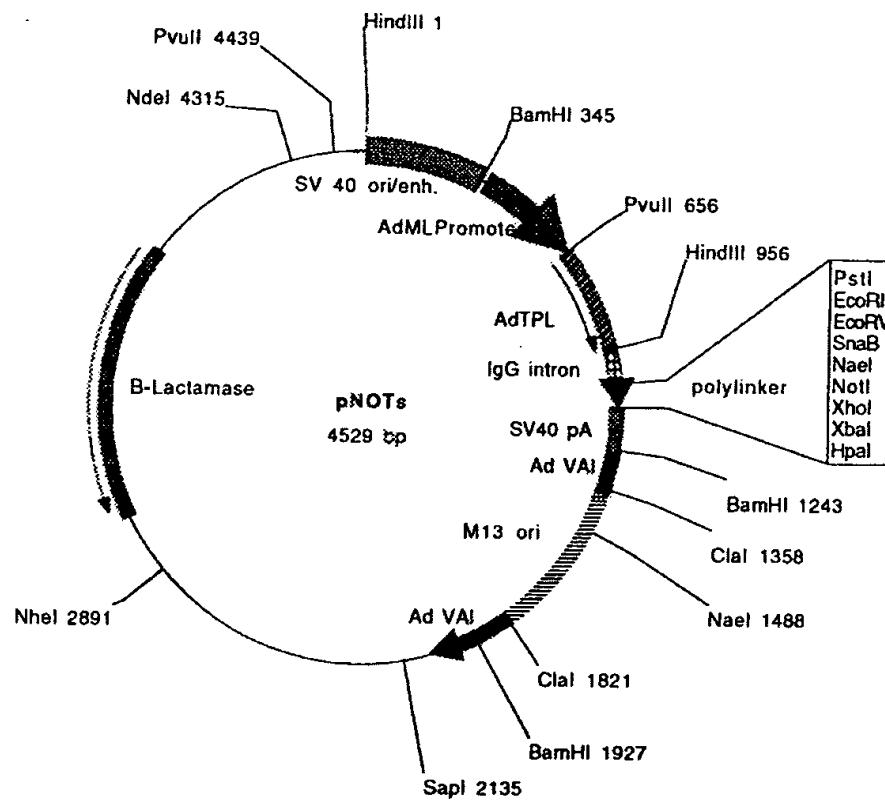
**Plasmid name:** pED6dpc2

**Plasmid size:** 5374 bp

**Comments/References:** pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.



FIGURE 1B



**Plasmid name:** pNOTs

**Plasmid size:** 4529 bp

**Comments/References:** pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

## SEQUENCE LISTING

<110> Jacobs, Kenneth  
 McCoy, John M.  
 LaVallie, Edward R.  
 Collins-Racie, Lisa A.  
 Evans, Cheryl  
 Merberg, David  
 Treacy, Maurice  
 Agostino, Michael J.  
 Steininger II, Robert J.  
 Genetics Institute, Inc.

<120> SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

<130> GI 6067A

<140>

<141>

<160> 32

<170> PatentIn Ver. 2.0

<210> 1

<211> 2360

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (2343)

<400> 1

```

ccacttcccc cttttgttaa ttaaaactaa gaagtcggaa tgggaacgag gtgcccagct 60
cccggtggaga aagctttaagg acaccacgcc agtgctttcc tgccttcctt ccgagatgga 120
aagaggagct cctagctcac ttaagccggg gtagggctgg ttctcctttc cgagccaaaa 180
tcccaggcga tgggtgaatta tgaacgtgcc acaccatgaa gctcttgtgg caggtaaactg 240
tgcaccacca cacctggaat gccatcctgc tcccgttcgt ctacctcacg gcgcaagtgt 300
ggattctgtg tgcagccatc gctgctgccg cctcagccgg gcccagaac tgcccctccg 360
tctgctcgtg cagtaaccag ttcagcaagg tgggtgtcac gcgccggggc ctctccgagg 420
tcccgcaggg tattccctcg aacacccggg acctcaacct catggagaac aacatccaga 480
tgatccaggc cgacaccttc cgccacctcc accacctgga ggtcctgcag ttgggcagga 540
actccatccg gcagattgag gtgggggcct tcaacggcct ggccagcctc aacacctgg 600
agctgttcga caactggctg acagtcaccc ctacgagggc ctttgaatac ctgtccaagc 660
tgcgggagct ctggcttcgc aacaacccca tcaaaagcat cccctcttac gccttcaacc 720
gggtgcccct cctcatgcgc ctggacttgg gggagctcaa gaagctggag tatatctctg 780
agggagcttt tgaggggctg ttcaacctca agtatctgaa cttgggcatg tgcaacatta 840
aagacatgcc caatctcacc cccctggtgg ggctggagga gctggagatg tcaggggaacc 900
acttccctga gatcaggcct ggctccttcc atggcctgag ctccctcaag aagctctggg 960
tcatgaactc acaggtcagc ctgattgagc ggaatgcttt tgacgggctg gcttcacttg 1020
tggaactcaa cttggcccac aataacctct cttctttgcc ccatgacctc tttaccccg 1080
tgaggtacct ggtggagttg catctacacc acaaccttg gaactgtgat tgtgacattc 1140
tgtggtagc ctggtggcct cgagagtata taccaccaa ttccacctgc tgtggccgct 1200
gtcatgctcc catgcacatg cgaggccgct acctcgtgga ggtggaccag gcctccttcc 1260
agtgtctctc ccccttcac atggacgcac ctcgagacct caacatttct gagggctcga 1320
tggcagaact taagtgtcgg actcccceta tgtcctcctg gaagtgggtg ctgcccgaatg 1380
ggacagtgtc cagccacgcc tcccgcacc caaggatctc tgtcctcaac gacggcacct 1440
tgaacttttc ccacgtgctg ctttcagaca ctggggtgta cacatgcatg gtgaccaatg 1500
ttgcaggcaa ctccaacgcc tcggcctacc tcaatgtgag cacggctgag cttaacacct 1560

```

```

ccaactacag cttcttcacc acagtaacag tggagaccac ggagatctcg cctgaggaca 1620
caacgcgaaa gtacaagcct gttcctacca cgtccactgg ttaccagccg gcatatacca 1680
cctctaccac ggtgtcatt cagactaccc gtgtgcccac gcaggtggca gtaccgcga 1740
cagacaccac tgacaagatg cagaccagcc tggatgaagt catgaagacc accaagatca 1800
tcattggctg ctttgtggca gtgattgtgc tagctgccgc catgttgatt gtcttctata 1860
aacttcgtaa gcggcaccag cagcggagta cagtcacagc cgcccggact gttgagataa 1920
tccaggtgga cgaagacatc ccagcagcaa catccgcagc agcaacagca gctccgtccg 1980
gtgtatcagg tgagggggca gtagtgctgc ccacaattca tgaccatatt aactacaaca 2040
cctacaaacc agcacatggg gccactgga cagaaaacag cctggggaac tctctgcacc 2100
ccacagtcac cactatctct gaaccttata taattcagac ccataccaag gacaaggtag 2160
aggaaactca aatatgactc cttcccccaa aaaacttata aaatgcaata gaatgcacac 2220
aaagacagca acttttgtac agagtgggga gagacttttt cttgtatatg cttatatatt 2280
aagtctatgg gctggtttaa aaaaacagat tatattaaaa tttaaagaca aaaagtcaaa 2340
acnaaaaaaa aaaaaaaaaa
2360

```

<210> 2  
 <211> 653  
 <212> PRT  
 <213> Homo sapiens

```

<400> 2
Met Lys Leu Leu Trp Gln Val Thr Val His His His Thr Trp Asn Ala
  1              5              10              15

Ile Leu Leu Pro Phe Val Tyr Leu Thr Ala Gln Val Trp Ile Leu Cys
      20              25              30

Ala Ala Ile Ala Ala Ala Ala Ser Ala Gly Pro Gln Asn Cys Pro Ser
      35              40              45

Val Cys Ser Cys Ser Asn Gln Phe Ser Lys Val Val Cys Thr Arg Arg
      50              55              60

Gly Leu Ser Glu Val Pro Gln Gly Ile Pro Ser Asn Thr Arg Tyr Leu
      65              70              75              80

Asn Leu Met Glu Asn Asn Ile Gln Met Ile Gln Ala Asp Thr Phe Arg
      85              90              95

His Leu His His Leu Glu Val Leu Gln Leu Gly Arg Asn Ser Ile Arg
      100             105             110

Gln Ile Glu Val Gly Ala Phe Asn Gly Leu Ala Ser Leu Asn Thr Leu
      115             120             125

Glu Leu Phe Asp Asn Trp Leu Thr Val Ile Pro Ser Gly Ala Phe Glu
      130             135             140

Tyr Leu Ser Lys Leu Arg Glu Leu Trp Leu Arg Asn Asn Pro Ile Glu
      145             150             155             160

Ser Ile Pro Ser Tyr Ala Phe Asn Arg Val Pro Ser Leu Met Arg Leu
      165             170             175

Asp Leu Gly Glu Leu Lys Lys Leu Glu Tyr Ile Ser Gly Gly Ala Phe
      180             185             190

Glu Gly Leu Phe Asn Leu Lys Tyr Leu Asn Leu Gly Met Cys Asn Ile
      195             200             205

```

Lys Asp Met Pro Asn Leu Thr Pro Leu Val Gly Leu Glu Glu Leu Glu  
 210 215 220  
 Met Ser Gly Asn His Phe Pro Glu Ile Arg Pro Gly Ser Phe His Gly  
 225 230 235 240  
 Leu Ser Ser Leu Lys Lys Leu Trp Val Met Asn Ser Gln Val Ser Leu  
 245 250 255  
 Ile Glu Arg Asn Ala Phe Asp Gly Leu Ala Ser Leu Val Glu Leu Asn  
 260 265 270  
 Leu Ala His Asn Asn Leu Ser Ser Leu Pro His Asp Leu Phe Thr Pro  
 275 280 285  
 Leu Arg Tyr Leu Val Glu Leu His Leu His His Asn Pro Trp Asn Cys  
 290 295 300  
 Asp Cys Asp Ile Leu Trp Leu Ala Trp Trp Leu Arg Glu Tyr Ile Pro  
 305 310 315 320  
 Thr Asn Ser Thr Cys Gly Arg Cys His Ala Pro Met His Met Arg  
 325 330 335  
 Gly Arg Tyr Leu Val Glu Val Asp Gln Ala Ser Phe Gln Cys Ser Ala  
 340 345 350  
 Pro Phe Ile Met Asp Ala Pro Arg Asp Leu Asn Ile Ser Glu Gly Arg  
 355 360 365  
 Met Ala Glu Leu Lys Cys Arg Thr Pro Pro Met Ser Ser Val Lys Trp  
 370 375 380  
 Leu Leu Pro Asn Gly Thr Val Leu Ser His Ala Ser Arg His Pro Arg  
 385 390 395 400  
 Ile Ser Val Leu Asn Asp Gly Thr Leu Asn Phe Ser His Val Leu Leu  
 405 410 415  
 Ser Asp Thr Gly Val Tyr Thr Cys Met Val Thr Asn Val Ala Gly Asn  
 420 425 430  
 Ser Asn Ala Ser Ala Tyr Leu Asn Val Ser Thr Ala Glu Leu Asn Thr  
 435 440 445  
 Ser Asn Tyr Ser Phe Phe Thr Thr Val Thr Val Glu Thr Thr Glu Ile  
 450 455 460  
 Ser Pro Glu Asp Thr Thr Arg Lys Tyr Lys Pro Val Pro Thr Thr Ser  
 465 470 475 480  
 Thr Gly Tyr Gln Pro Ala Tyr Thr Thr Ser Thr Thr Val Leu Ile Gln  
 485 490 495  
 Thr Thr Arg Val Pro Lys Gln Val Ala Val Pro Ala Thr Asp Thr Thr  
 500 505 510  
 Asp Lys Met Gln Thr Ser Leu Asp Glu Val Met Lys Thr Thr Lys Ile  
 515 520 525

Ile Ile Gly Cys Phe Val Ala Val Ile Val Leu Ala Ala Ala Met Leu  
 530 535 540  
 Ile Val Phe Tyr Lys Leu Arg Lys Arg His Gln Gln Arg Ser Thr Val  
 545 550 555 560  
 Thr Ala Ala Arg Thr Val Glu Ile Ile Gln Val Asp Glu Asp Ile Pro  
 565 570 575  
 Ala Ala Thr Ser Ala Ala Ala Thr Ala Ala Pro Ser Gly Val Ser Gly  
 580 585 590  
 Glu Gly Ala Val Val Leu Pro Thr Ile His Asp His Ile Asn Tyr Asn  
 595 600 605  
 Thr Tyr Lys Pro Ala His Gly Ala His Trp Thr Glu Asn Ser Leu Gly  
 610 615 620  
 Asn Ser Leu His Pro Thr Val Thr Thr Ile Ser Glu Pro Tyr Ile Ile  
 625 630 635 640  
 Gln Thr His Thr Lys Asp Lys Val Gln Glu Thr Gln Ile  
 645 650

<210> 3  
 <211> 1558  
 <212> DNA  
 <213> Homo sapiens

<400> 3  
 gcgaagctga taagccaata agagtgtgt ctagaacgca ttgcaaacc gttaagagta 60  
 gcttcttttt gaggcctgga cttctgattc atctatagaa gtctattatg attccatcaa 120  
 ttgcatagct tgtgtacctg gtagaaattg tgtcttgga tgacctttc gagttattga 180  
 catggctctg atgaatagaa catgagcccc aaaactaaat ccaaaggaa ttttctatct 240  
 ttcatccccc acatgtggca agacaagttg gccctttctt acccagaggt cttttgtgtg 300  
 actgcatctt tctcctcctg tctccattgt gtgctttcca ttttgtctt agtcctata 360  
 ctgttaggtg ttttcttcac tggcattcac aaatttaagc cattgtgcc tcattagcct 420  
 tgtattttgt gtgcataatca tatatccaga cctgtatgtt cactttaagc attcttata 480  
 cacactgtct cctcatctac catatggtaa atgaggagac agtgtcttat atcacactgt 540  
 ctctcatctt accatatggt aaatgttaaa actccacatt tgtctgcac agggaaaatg 600  
 catgggcaca catcctcctt cctcctctt ctgctctcct cccttcttc aggcctctta 660  
 gcattgtttg ttttccatt tctgatacta ctactccatg ctgaagattt gccatattac 720  
 tattttggaa acattgagtg atagaactcc tagaaaattt gcaaagaaat gttacatact 780  
 gtatatcaaa ctctcagatt ctagtgttga aaaagtagcc tatactttgc tattacttat 840  
 acctgctgcc atagaaaaaa ataagtttat tcatgacaca ttacatttg atcataaata 900  
 aaagaaaaaa gggcaccttt ttggagttag tcatggtagt cattagtgat atttctgaac 960  
 agttcttaat ttaaaatact tcaaaggaag taaagggtcat ggcttagctg aaggaaatgc 1020  
 tccagaaatt ggactgtgta aaccatcagt acaataatac gctgtgtatg tatgtgtata 1080  
 taaaatgaga attatggcat aattggagca ttgcatata tcaacaaact cacattgaga 1140  
 caaaacttag ttttacagct gtcttgatta aagccaagtg ttccatgttg ctgtgaagaa 1200  
 tagcctcttt caaatacttt ggaagtagt tacttgaaa ctgttaaagg tattacattt 1260  
 ttatatttaa acacctatag agatcttcaa ttcttgagt ctgagcttg ggggtgaatt 1320  
 ctaaatgtgt atcataatct gtctttgtg aaacattttg aaaatatgta tatataatat 1380  
 tgtatatgca aattgtgttg ttctacttgt aaagggaaaa ggcttatttt tctttatatt 1440  
 tctgataact tgttttgc atgaccagca ctgactgaaa ggcatgtgta gctgcaaaaa 1500  
 ctgttgcttt ttttgtgaaa tgaataaaa agtattttaa tacaacaaaa aaaaaaaa 1558

<210> 4  
 <211> 71

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

Met Ser Pro Lys Thr Lys Ser Lys Arg Asn Phe Leu Ser Phe Ile Pro  
 1 5 10 15

His Met Trp Gln Asp Lys Leu Ala Leu Ser Tyr Pro Glu Val Phe Cys  
 20 25 30

Val Thr Ala Ser Phe Ser Ser Val Leu His Cys Val Leu Ser Ile Leu  
 35 40 45

Ser Leu Val Pro Ile Leu Leu Gly Val Phe Phe Thr Gly Ile His Lys  
 50 55 60

Phe Lys Pro Leu Leu Pro His  
 65 70

&lt;210&gt; 5

&lt;211&gt; 2637

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 5

ccttgagcct gttcctttac ctcaccaggc ctcattttcc acatctgttg aacaatctgg 60  
 ctcttagggc ccttttcaat tctagtgttc tggagggtcca tgaactgttt ttgtgtagta 120  
 gagaagaact ttcctgaagt tcgaatataat aatatatgac agaggaaatc actgttatga 180  
 accttagagt tattcccatg ggaggatgaa gactgacatt gcttccttgc acaccaggac 240  
 ttgggcctgt actgctctaa ttctgaactt aatttcctgt gtggctcttg cttgggtcat 300  
 gggacggtct cctaacataa tgtttatata ccgtggctca ggcatagaga cccactcatg 360  
 gtttcccatc agaatttggg ttccagttca ccactccaca tacacacacc aaggctaagt 420  
 attcatgcta ttgtcagttt ttgcttcttt gatattgttc tgaaggtagt ttttattgat 480  
 acttctttct ttctaaagtc aatgagtaaa gttatggcaa gacattggaa agacattgca 540  
 gtgctccag gctggagtgc atgactcatc tcctggcccc tccaggaaac cggagaaggc 600  
 agaaaagggc aagaatgttg tgttgtcttt gggctgtgga gttcatgatt atggacactt 660  
 tctgaattat tggcactaag ggcacatcct gcttagcatg gaggctggac agctctggga 720  
 aagcagagat tctttttagg gagtccagag catttttttc ttctttctta tccccctctt 780  
 ccaatagaga tgtcagaagg gtgactgtta gggttgatac attatgtggt gctctgtttt 840  
 tttagacatt gtctactcag agtgattacc aaggtggagc tcattatgtt acctggcacc 900  
 tgcaactact tgccctcatgt tggactaagg tggaaagact cctattggga attcactcag 960  
 ggaggcacca ggcattgtta gaaaagcagt atcgagagc acttaagagc tgagccttgt 1020  
 gatcgggttg acttgggttt atgtcctcgc actgtcgcct cctagattta tgacctcagt 1080  
 ttctctcatt ctaaagtaga gacagggatg gagggtgggg agaatatgaa ttaatttggga 1140  
 ttaaatggat gagaatcaga agtatttatt tgggaattggt gtgaattgga gcacatgggg 1200  
 aactgaaaac aaagggtgata atttatgttg ttagatgaa ctaagaataa tgaaaacagc 1260  
 atgaaggatg gaatcaaatt aagtgcacac atataaagga aatggactgg taaaagctcg 1320  
 agggcaaaag tgtatggaga tgggaattca ggttggaaag aggctgctga tacaaggact 1380  
 ggagttggga gagctgctga ccaaatgttg aaatgaagat tttttaaaaa tgggcgctgg 1440  
 ttggattcag gcaagctatt ttaaccagg actcacttta attcagtcaa aatatttctt 1500  
 atcttccagt tggggagggtg ttttgggaca ggaaaaaaag ggagataaga ataccatccc 1560  
 agagaaagaa gagaagtatt tgcattttgt tctctagagt tgcagatata tgctgtatct 1620  
 cttctatatg gtaatacatt gcaaaggaga tacagtgaag aaatgaagta aaataaaatg 1680  
 tgtgtgtctt cctctcatat gcataaagga agacaatata tgttggtaac atggatttgc 1740  
 ttgattgaa gaatttttaa ggtgttatcc atttaattct tatatagaca gaatttactg 1800  
 agctgaaaaa gatatgcatt tttttccttc tgtggctgtt ttttaggat gctgaatcct 1860  
 tcaaaatttt tgccctggag gatattaggc aatgaggtga ttttttattt attatttttt 1920  
 aaggtagcag atgttcacat tggtagctag ggcacaaacg cagggtgccc aagggaagtaa 1980  
 agtttagagat gggggtactg agctggagggt gcaggataac gaggcgagga atgggacaaa 2040

```

gggaatatta tgaagtagca ggatttagtg aaaggaagca tataaaaatg gacataaggg 2100
ccagggtgagg tggctcacgc ccctaatect agcacttttg gaggcctagg cagggtggatc 2160
acctaagggtc aggagttcga gaccagcctg accaacaccc gtctctacta aaagtacaaa 2220
aattagctgg gcgtgtggcg cgtgcctgta gtcccagcta ctcaggaggc tgaggcagga 2280
gaatcgcttg accccaggag gtggagggtg cagtgcgccg agattgcgtc actgcactcc 2340
agcctgggtg acagagttag acaccgtctc aaaacaata aaaattaaaa aaaatggaca 2400
tatggtttat tcattgggaa gagcacctag tgttggtatg agtgtaaaaa tgttcttagc 2460
agccaagtca ttgcagacct ttctctccta ctmtgtgttc tggctctcat atatttttct 2520
ttctcacaaa ttctaccatt atgatcatag ttgtattcat cttcttgagt cagctaaata 2580
atccttatat tttgatgaaa ggtaaccttc acttccaata ttaaaaaaaaa aaaaaaa 2637

```

<210> 6  
 <211> 70  
 <212> PRT  
 <213> Homo sapiens

```

<400> 6
Met Lys Thr Asp Ile Ala Ser Leu His Thr Arg Thr Trp Ala Cys Thr
  1           5           10          15
Ala Leu Ile Leu Asn Leu Ile Ser Cys Val Ala Leu Ala Trp Phe Met
          20          25          30
Gly Arg Ser Pro Asn Ile Met Phe Ile Ser Arg Gly Ser Gly Ile Glu
          35          40          45
Thr His Ser Trp Phe Pro Ile Arg Ile Trp Val Pro Val His His Ser
          50          55          60
Thr Tyr Thr His Gln Gly
          65          70

```

<210> 7  
 <211> 4398  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> unsure  
 <222> (4338)

```

<400> 7
ggcgctggtg gctgcgggcg cggcgggcggc agcggcgctc gagcgggttc tgtcagggtc 60
agccggcggg cccctggtt ggtccacctg caaatcgcg agcggcgccc cagggtatcga 120
tggcgatgaa ctataacgcg aaggatgaag tggacggttg gccccgtgt gctccggggg 180
gcaccgcgaa gactcggaga cgggataaca cggccttcaa acagcaacgg ctgccagctt 240
ggcagcccat ccttacggct ggcacgggtg tactatttt cttcatcatc ggtctcatct 300
tcattcccat cggcattggc atttttgtca cctccaacaa catccgcgag atcgagattg 360
attataccgg aacagagcct tccagtcctt gtaataaatg tttatctccg gatgtgacac 420
cttgcttttg taccattaac ttcacactgg aaaagtcatt tgaggggcaac gtgtttatgt 480
attatggact gtctaatttc tatcaaaacc atcgctgtta cgtgaaatct cgagatgata 540
gtcaactaaa tggagattct agtgctttgc ttaatcccag taaggaaatg gaaccttacc 600
gaagaaatga agacaaacca attgctcctt gtggagctat tgccaacagc atgtttaatg 660
atacattaga attgtttctc attggcaatg attcttatcc tatacctatc gctttgaaaa 720
agaaaggtat tgcttggttg acagataaaa atgtgaaatt cagaaatccc cctggaggag 780
acaacctgga agaacgattt aaaggtacaa caaagcctgt gaactggctt aaaccagttt 840
acatgctgga ttctgaccca gataataatg gattcataaa tgaggatttt attgtttgga 900
tgcgactgac agcattacct acttttcgca agttgtatcg tcttatagaa aggaaaagtg 960
atttacatcc aacattacca gctggccgat actctttgaa tgtcacatac aattaccctg 1020

```

```

tacattattt  tgatggacga  aaacggatga  tcttgagcac  tatttcatgg  atgggaggaa  1080
aaaatccatt  tttggggatt  gcttacatcg  ctgttggtac  catctccttc  cttctgggag  1140
ttgtactgct  agtaattaat  cataaatata  gaaacagtag  taatacagct  gacattacca  1200
tttaatttta  tattatgaaa  gcaaatcatc  tgcattgtgca  tcaaggccag  tcctattcaa  1260
cctagctttc  gaatgctgat  atctgggttag  tatgtcattt  tgaagttggc  acataacttt  1320
tctaaaaaaa  agcagctctt  gttgtttgct  tcttccctac  ggatgacttc  taaaaatata  1380
tgacgggtat  aaaaaaatta  gctatattga  tcatatcaac  actgtaactg  ctgaaatggc  1440
atttctaatgt  ttgcttttta  ttcggacagg  ccacatgatg  catagagcct  ctttcatgtg  1500
acttgtgtct  actgcttaaa  tctttatgct  gtgttgatga  tattatattg  acatatgaag  1560
ctgtatatgt  gtatgtattt  tgtggagaag  agggattaca  agatgtatga  gtataatgac  1620
ttgctaacct  ttcaggattc  cgagaaagat  gaagaaagac  catatctaaa  taatacactt  1680
catcattttc  atgtgtataa  atgcttaaag  taccatcttt  gttgaggtgg  ttcattgtatc  1740
cagtttatcc  agtacagtta  tttgtcaagc  tttagctttga  tttcaaagga  cacgcttacc  1800
ttgtctggca  taagaattaa  tgctcatgtc  tgcagtggtt  gggtaggtcc  tgcttaggag  1860
aattaaaaaa  ttcctctttc  cgtttggttg  aatgttgcat  tcaggaaccc  caactcactt  1920
ggaatgtttt  tatatgtaat  catttccctt  gaagcttata  ctttataagg  gaagaaagaa  1980
ttcagtggtg  atgggaaaac  tgcttggcag  accttcatct  tctgcctcaa  ctgtaaacca  2040
catgtaaatg  cttaatggag  actgttttca  ttcttgtgat  atttaacatt  cagaaaatta  2100
cttcagcttt  ggaaatactc  aggcgtgttt  tattctgcag  gtaagtgttt  tgacttaagt  2160
actaatattc  cagaaatttt  tgaagcagc  aaccttaatt  tcctatgtat  ttcattccac  2220
ttttgcatat  aggtcaaata  gcaatgtgta  tgcacattct  ctttagttaa  ggcaccaatt  2280
gttttgggtg  gttttcctaa  gacatacttt  aaaaagatgt  tctataaatt  tcctagttaa  2340
attatgggga  ttttggagta  tgtacatgat  aaattataat  acgtatatgg  ttgaagttat  2400
tttattttt  actaatgaat  tattttaata  ttctttattg  aataaatgct  gtaacttggt  2460
tgctatggaa  cttattctta  aagttctagt  taaaaataat  ttttccacat  gcatgaaaat  2520
atgtattaat  cagaggtggc  ttaattacat  tgaattgtct  tttttgtgt  tgttttttta  2580
ctgaaataac  tcatgtttgt  gtagaagaat  gcctgtttac  tcagagttaa  tattttcctt  2640
cagttatatt  ttaaatcaaa  aggtctgggt  aatgtatact  tttgattaat  atatactttt  2700
tttaaaaaaa  aaaaaacaat  gtaatggtta  atagtagaaa  tgtgccacac  ttttcaagtt  2760
ttatataaca  tattgaaatt  agttaaaaga  atgtgtgttt  cataatgact  tttactgggt  2820
aaaaatatta  cttgcacgaa  gtacttgatg  tatggttatc  ctgaaatttc  ggagtatttg  2880
gtgtgttctt  tgtctaaaaa  tagtctgttt  tgtcagtcct  tcagaatatt  atttattctg  2940
aagattgtcc  ctcttgcact  tggcagttta  ttttcgggga  tacattgttg  ggggagaggg  3000
gtttctgcca  ctctttccag  attgagcttg  tgctgtttta  ggaggactac  catcctgcaa  3060
ctctttttct  aattggggca  cagaggatgt  cgctaaagaa  aagttgaaga  gccctttcag  3120
cactttctca  tctgtggaga  agatggaaat  ttaaaataca  tttggagtgt  tatctgtttt  3180
acaagtccat  tgatggccta  agttcctcct  gttttctgct  gtttgatctc  taagggaactc  3240
ctgttgctaa  atatgaagag  tatggcacat  tcatatagtc  tctgtgaagc  atggggggag  3300
ggaagacatt  tctttttctt  ataggcttta  tgctcaaatg  tcatagtctc  ctttcaaaga  3360
atttgtgtgc  attttaaatg  caccagctt  aagtagaaga  cattgaagga  tgcatttaatt  3420
ttcaggaact  attttgaatt  atgaaaagat  tcccaattga  aaaaattatt  caacaagtaa  3480
aagctaagaa  atttcattga  aatcataagg  cagtttaagc  ataaattgat  aaaaatagct  3540
gtgtactact  aattaataga  aaatcattca  accaagagaa  gagtcaagtg  aatatcgttt  3600
gtttatttgc  tagtgagttt  ctttgaacg  ttgattttat  taaatgataa  tatttggtta  3660
gtatgtccta  tgtaataaaa  aatgaacaaa  attaattttg  ctatgttcag  gtgtcttgat  3720
aaaaatacaa  tgcctcagtg  ttgttgctta  catttagcac  taaattttta  cacagggtca  3780
gtgagtcag  gtttttaact  cttcatgcct  ggatgggata  aaatgtaatt  cattgttaa  3840
ttaattcata  tttgtattta  ttaatcactg  tgacaacatt  aaccatttgt  tcttaccagg  3900
aagtggtcag  attatcatct  gagttacagt  tagactggct  aagtttggt  ttagatcaag  3960
gggaatgtcc  agtaaacaga  gaggtaaagc  tgatggaaat  aatgaagtgg  ggtacacagg  4020
aaaaacctga  ctagttagga  ggagcagctg  agagataggg  tcagtgaatg  cggttcagcc  4080
tgetacctct  cctgtcttca  tagaaccatk  gccttagaat  tattgtatga  cacgtttttt  4140
gttggttaag  ctgtaaggtt  ttgttctttg  tgaacatggg  tattttgagg  ggagggtgga  4200
gggagtaggg  aagtggctct  tttacaagaa  ttttgatgca  taagtgtcta  ttgtagggtt  4260
tggatgatct  agtaaagtgt  tttagaaccc  ctttttatcc  catgcacct  tcagtaaaaa  4320
taaaaatcac  aattctgnta  atgtcatttg  gaacttcaaa  ataaatatct  tgtctaaaaa  4380
caaaaaaaaa  aaaaaaaaaa  4398

```

&lt;210&gt; 8

&lt;211&gt; 361



&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 8

```

Met Ala Met Asn Tyr Asn Ala Lys Asp Glu Val Asp Gly Gly Pro Pro
 1           5           10          15

Cys Ala Pro Gly Gly Thr Ala Lys Thr Arg Arg Pro Asp Asn Thr Ala
      20           25           30

Phe Lys Gln Gln Arg Leu Pro Ala Trp Gln Pro Ile Leu Thr Ala Gly
      35           40           45

Thr Val Leu Pro Ile Phe Phe Ile Ile Gly Leu Ile Phe Ile Pro Ile
      50           55           60

Gly Ile Gly Ile Phe Val Thr Ser Asn Asn Ile Arg Glu Ile Glu Ile
      65           70           75           80

Asp Tyr Thr Gly Thr Glu Pro Ser Ser Pro Cys Asn Lys Cys Leu Ser
      85           90           95

Pro Asp Val Thr Pro Cys Phe Cys Thr Ile Asn Phe Thr Leu Glu Lys
      100          105          110

Ser Phe Glu Gly Asn Val Phe Met Tyr Tyr Gly Leu Ser Asn Phe Tyr
      115          120          125

Gln Asn His Arg Arg Tyr Val Lys Ser Arg Asp Asp Ser Gln Leu Asn
      130          135          140

Gly Asp Ser Ser Ala Leu Leu Asn Pro Ser Lys Glu Cys Glu Pro Tyr
      145          150          155          160

Arg Arg Asn Glu Asp Lys Pro Ile Ala Pro Cys Gly Ala Ile Ala Asn
      165          170          175

Ser Met Phe Asn Asp Thr Leu Glu Leu Phe Leu Ile Gly Asn Asp Ser
      180          185          190

Tyr Pro Ile Pro Ile Ala Leu Lys Lys Lys Gly Ile Ala Trp Trp Thr
      195          200          205

Asp Lys Asn Val Lys Phe Arg Asn Pro Pro Gly Gly Asp Asn Leu Glu
      210          215          220

Glu Arg Phe Lys Gly Thr Thr Lys Pro Val Asn Trp Leu Lys Pro Val
      225          230          235          240

Tyr Met Leu Asp Ser Asp Pro Asp Asn Asn Gly Phe Ile Asn Glu Asp
      245          250          255

Phe Ile Val Trp Met Arg Thr Ala Ala Leu Pro Thr Phe Arg Lys Leu
      260          265          270

Tyr Arg Leu Ile Glu Arg Lys Ser Asp Leu His Pro Thr Leu Pro Ala
      275          280          285

Gly Arg Tyr Ser Leu Asn Val Thr Tyr Asn Tyr Pro Val His Tyr Phe
      290          295          300

```

Asp Gly Arg Lys Arg Met Ile Leu Ser Thr Ile Ser Trp Met Gly Gly  
305 310 315 320

Lys Asn Pro Phe Leu Gly Ile Ala Tyr Ile Ala Val Gly Ser Ile Ser  
325 330 335

Phe Leu Leu Gly Val Val Leu Leu Val Ile Asn His Lys Tyr Arg Asn  
340 345 350

Ser Ser Asn Thr Ala Asp Ile Thr Ile  
355 360

<210> 9  
<211> 777  
<212> DNA  
<213> Homo sapiens

<400> 9  
agaaattata gaggtagtgt attttccttg tgtatatatt aaacacaccc atacaaacat 60  
tggtaaagtt gattacatca aagaatcttt agcttatctt ttgaagctac tggatattat 120  
tggtctctct aggtttttat ataaatagtg aaatctgaat tactgaaaac catgttaatt 180  
tttagaactc attttcctca gtagagacta gtgatgcatt agcttctggg aacaaacttg 240  
tatcggttct taattaaatt atccaaaacg gaggcattta aacacttgga ttacaccag 300  
tcttttggtg ttgcttttta aaataaagtg ctctgtattg tattctccat attttggagt 360  
aattatctac atgatgttta tagttcctgt ggtttttcac ccaagaagca gaatctcatt 420  
cagtacattt agttttataa gagtcacgaa gctaaatcct tgggctatgt cagaggcaca 480  
aagtctagaa tgtgtgtatt cacaatgggtg tatgtacatt ttgtgccttg attcacttag 540  
aagtgtctca gaaaacctgg acagttcgct tctacacaag aattttatat gtatttatga 600  
agatgattct gtaccctagt atatcttttt gggcatggac taatttgat ctgtttaact 660  
catattctgc acgatctgta tatagtacat caaacttaga ggtgtgacct taaatttaac 720  
tttttttaaa aactggggagg tcaataaaat ttaaactgct taataaaaaa aaaaaa 777

<210> 10  
<211> 36  
<212> PRT  
<213> Homo sapiens

<400> 10  
Met Lys Met Ile Leu Tyr Pro Ser Ile Ser Phe Trp Ala Trp Thr Asn  
1 5 10 15  
Leu Tyr Leu Phe Asn Ser Tyr Ser Ala Arg Ser Val Tyr Ser Thr Ser  
20 25 30  
Asn Leu Glu Val  
35

<210> 11  
<211> 1378  
<212> DNA  
<213> Homo sapiens

<400> 11  
ctccgctggc aacggcgccg ctccccgctc ctccctccca gccatggcgt tcacgttcgc 60  
ggccttctgc tacatgctgg cgtgctgct cactgcccgc ctcatcttct tcgccatttg 120  
gcacattata gcatttgatg agctgaagac tgattacaag aatcctatag accagtgtaa 180  
taccctgaat ccccttgtag tcccagagta cctcatccac gctttcttct gtgtcatgtt 240

```

tctttgtgca gcagagtggc ttacactggg tctcaatatg cccctcttgg catatcatat 300
ttggagggtat atgagtagac cagtgatgag tggcccagga ctctatgacc ctacaacccat 360
catgaatgca gatattcttag catattgtca gaaggaagga tgggtgcaaat tagcttttta 420
tcttctagca tttttttact acctatatgg catgatctat gttttggtga gctcttagaa 480
caacacacag aagaattggt ccagttaagt gcatgcaaaa agccaccaa tgaagggatt 540
ctatccagca agatcctgtc caagagtagc ctgtggaatc tgatcagtta ctttaaaaaa 600
tgactcctta ttttttaaat gtttccacat ttttgcttgt ggaaagactg ttttcatatg 660
ttatactcag ataaagattt taaatggtat tacgtataaa ttaataataa atgattacct 720
ctgggtgttg caggtttgaa cttgcacttc ttaaggaaca gccataatcc tctgaatgat 780
gcattaatta ctgactgtcc tagtacattg gaagcttttg tttataggaa cttgtagggc 840
tcattttggt ttcattgaaa cagtatctaa ttataaatta gctgtagata tcaggtgctt 900
ctgatgaagt gaaaatgtat atctgactag tgggaaactt catgggtttc ctcactctgc 960
atgtcgatga ttatatatgg atacatttac aaaaataaaa agcgggaatt tcccttcgc 1020
ttgaatatta tccctgtata ttgcatgaat gagagatttc ccataatttc atcagagtaa 1080
taaataact tgctttaatt cttaagcata agtaaaccatg atataaaaa atatgctgaa 1140
ttacttgatga agaatgcatt taaagctatt ttaaatgtgt ttttatttgt aagacattac 1200
ttattaagaa attggttatt atgcttactg ttctaactctg gtggtaaagg tattcttaag 1260
aatttcagg tactacagat tttcaaaact gaatgagaga aaattgtata accatcctgc 1320
tgttccttta gtgcaataca ataaaactct gaaattaaga ctcaaaaaa aaaaaaaa 1378

```

&lt;210&gt; 12

&lt;211&gt; 144

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 12

```

Met Ala Phe Thr Phe Ala Ala Phe Cys Tyr Met Leu Ala Leu Leu Leu
  1              5              10              15

Thr Ala Ala Leu Ile Phe Phe Ala Ile Trp His Ile Ile Ala Phe Asp
  20              25              30

Glu Leu Lys Thr Asp Tyr Lys Asn Pro Ile Asp Gln Cys Asn Thr Leu
  35              40              45

Asn Pro Leu Val Leu Pro Glu Tyr Leu Ile His Ala Phe Phe Cys Val
  50              55              60

Met Phe Leu Cys Ala Ala Glu Trp Leu Thr Leu Gly Leu Asn Met Pro
  65              70              75              80

Leu Leu Ala Tyr His Ile Trp Arg Tyr Met Ser Arg Pro Val Met Ser
  85              90              95

Gly Pro Gly Leu Tyr Asp Pro Thr Thr Ile Met Asn Ala Asp Ile Leu
 100              105              110

Ala Tyr Cys Gln Lys Glu Gly Trp Cys Lys Leu Ala Phe Tyr Leu Leu
 115              120              125

Ala Phe Phe Tyr Tyr Leu Tyr Gly Met Ile Tyr Val Leu Val Ser Ser
 130              135              140

```

&lt;210&gt; 13

&lt;211&gt; 1450

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 13

```

cgaaaatgtc aagagaaagg ccaaagattt gtactttttc acttacaaag cactcctttt 60
tcccttaaac ttctttctgt caaattagat ttaatgagag agtactattt ttaaggagct 120
atctgtttat gtagaatgat tttgttaaga gtaatgtaaa ctattattga gtagaggcct 180
aaagaggact gtgcattttt gctattttaa ggaatcacaa atgatacatc ttaagtgcgc 240
aaaaatgaca agttttacta gctaagtaga gaaataaata tcaaatgcag cgctacaatt 300
ttcattatct taagtacatt gtacatttct acagaacctg tgattattct cgcatagata 360
ggatgggtact tgcataatgg gaattactac tgttgacagt ttccgcagaa atcctatttc 420
agtggacca cttgttgcca tggcagcaaa tgccaacatt ttgtggaata gcagcaaatc 480
tacaagagac cctgggtggg ttttcgtttt gttttctttg ttttttcccc cttctcctga 540
atcagcaggg atggaaggag ggtagggaag ttatgaatta ctcttccag tagtagctct 600
gaagtgtcac atttaatatc agtttttttt aaacatgatt ctagttaaag gtagaagaga 660
gaagaaagag gaagtgttca cttttttaat acactgattt agaaatttga tgtcttatat 720
cagtagttct gaggtattga tagcttgctt tatttctgcc ttacgttga cagtgttgaa 780
gcagggtgaa taactagggc atatattttt tttttttttg taagctggtt catgatgttt 840
tctttggaat ttccggataa gtccaggaaa acattctgca tgttgatctt agtctgatgt 900
acttatccat ctcatcaaa acaaaaacac acagactgca ttgtagctc tgtaatcctt 960
gaatacggaa gtaaattttt ttctttcctg actttgacat tgtagctata ctgtttccat 1020
ttttgttttt acaaatcctt tgggtctaat tctgtgagcc tacctatagc actggattaa 1080
aatgtctgca tcatttcttt agttatccag ttaactttaa aactgttgta aaagtgtaaa 1140
ccagcccatg acaggttttt gwacatgtta aagaacttca ttgttcagtt ttcatgatta 1200
ttgtgtaagg aagactgatg tagatgttct gtgctgtcct ggaccatggt aattacactt 1260
acgacgtatt ttagttccac atcacaatga ttgtcccca gtgacccttt tatcctttct 1320
aggcacattt cttgttggtt ttgttggtgc agttccctt tgcattgtat tgctttgaca 1380
actgtaattt gaatcagatc tgaaagaggt ccagaataaa atatattttg atattaaaaa 1440
aaaaaaaaa                                     1450

```

&lt;210&gt; 14

&lt;211&gt; 102

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 14

```

Met Gln Arg Tyr Asn Phe His Tyr Leu Lys Tyr Ile Val His Phe Tyr
  1             5             10             15

```

```

Arg Thr Cys Asp Tyr Ser Arg Met Ile Arg Met Val Leu Ala Tyr Gly
      20             25             30

```

```

Glu Leu Leu Leu Leu Thr Val Ser Ala Glu Ile Leu Phe Gln Trp Thr
      35             40             45

```

```

Asn Ile Val Ala Trp Gln Gln Met Pro Thr Phe Cys Gly Ile Ala Ala
      50             55             60

```

```

Asn Leu Gln Glu Thr Leu Val Gly Phe Ser Phe Cys Phe Leu Cys Phe
      65             70             75             80

```

```

Phe Pro Leu Leu Leu Asn Gln Gln Gly Trp Lys Glu Gly Arg Glu Val
      85             90             95

```

```

Met Asn Tyr Ser Phe Gln
      100

```

&lt;210&gt; 15

&lt;211&gt; 3546

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 15

```

gtggaattgg acactgggtgc cccaagcgag gagttgagca gtgctggaga agtaacgaaa 60
cagacagtct tacmgaagga agaggagagg agtcagccaa ctaaaacccc ttcatcttct 120
caagagcccc ctgatgaagg aacctcaggg acagatgtga acaaaggatc mtcaaagaat 180
gctttgtcct cmatggatcc tgaagtgagg cttagtagcc ccccagggaa gccagaagat 240
tcatccagtg ttgatgggtca gtcagtgggg actccagttg ggccagaaac tggaggagag 300
aagaatgggc cmgaagaaga ggaagaagag gactttgatg acctcaccca agatgaggaa 360
gatgaaatgt catcagcttc tgaggaatct gtgctttctg tcccagaact ccaggaaacm 420
atggagaaac tgacttggct ggcactctga aggcgcagta gtcaggaggg tgagtctgaa 480
gaagagaatt ctcaggagga gaactctgag ccagaagaag aggaggaaga agaagcagaa 540
ggaatggaaa gcctgcagaa agaggatgaa atgacggatg aagcagttgg agactctgct 600
gagaagcctc ctacttttgc ttacactgag actgctccag aagtggagac cagcagaact 660
ccaccaggag agagcatcaa agctgctgga aaaggccgga acaatcatcg agctcgcaac 720
aagcggggaa gtcgggctcg ggccagcaag gacacctcca agctgctgtt gctgtatgat 780
gaggacattc tcgagcgaga tccactcagg gacagaagg acttggcctt tgcccaagct 840
tatctgacca ggggtcgaga agccctacaa catatccctg gcaagtatga agacttcctt 900
caagtcattc atgaatttga gtcaagtacc cagagacgga cggtgtaga tctctacaaa 960
agcctgcaaa ttctgctcca agactggcct cagctgttga aagactttgc tgctttcctg 1020
ttacctgagc aagctctggc ctgtggatta tttgaggagc agcaggcttt tgagaagagc 1080
cgcaagtacc ttcggcagct ggagatttgc tttgcagaga acccctcaca ccaccagaag 1140
attatcaagg tcctccaagg ctgtgcagac tgctttcccc aggagatcac cgagctcaag 1200
acacagatgt ggcagctcct caagggccac gaccacctgc aggatgagtt ttctatcttc 1260
tttgaccact tgcgcccagc agctagccgg atgggtgact ttgaagagat caattggact 1320
gaggaaaagg agtatgagtt tgatggcctt gaagaagtgg ccctgcctga tgtggaagaa 1380
gaggaggagc ctcccaagat acccacagcc tcaaagaaca agaggaaaa agagatcggg 1440
gtccaaaatc atgataagga gactgaatgg ccagatgggg ccaaggactg tgctgtctcc 1500
tgccatgaag gaggctcaga ttccaagctg aagaagagca aaaggcgag ctgtagccac 1560
tgtagcagca aggtctgtga cagcaaatcc tacaagagca aggagcccca tgagttggtg 1620
ggcagcagcc cccaccgaga ggctagtcct atgcctggtg ctaaggaagc tgggcagggc 1680
aaggatatga tggaaagaga agccccagag gagcgggaga gactgaggc caccagagc 1740
aggactgtca ggaccaccag aaaggagag atgcctgttt caggattggc agtggggagc 1800
actttgccat cccctcgaga agtgactgtt acagaacggc tcctcctgga tggccacca 1860
cctcattcac cagagactcc tcaatttccc ccacaactg gagctgtact gtacactgtt 1920
aagagaaacc aggttggggc tgaggttcgc tcctgcccc aggcattccc cagacttcag 1980
aaagagaggg agggccaaaa ggcatgagt gactcagagg ctttgatgct ggtctgggat 2040
gcatcagaaa ctgagaaatt gcctggtacc gtgggaaccc cctgcttctt tcctgagtc 2100
tgtttccctc aagaccagag atgcaggag aagacatgtg tccgggaaac cagacactca 2160
agagtgcatt cccctcgaag gcagagctcg ggtgaagaca agagacagga cgtgccctgt 2220
ccatgaatct ccatcaggaa ttgacacctc agagacttct cccaaagccc ctagaggggg 2280
tttggtctaa gacagtggaa cacaggccaa ggtccagag ggggagcagc agccaaaggc 2340
cgagaagct acggtgtgtg ccaacaacag caaggctcag tccactgggg aaaaggttgt 2400
cctgtggaca agggaagctg accgtgtgat cctcaccatg tgccaggagc aaggggcaca 2460
gccacagacc ttcaacatca tctcccagca gctgggaaat aagaccctg ctgaggttcc 2520
ccaccgtttt cgagaactca tgagctctt ccacactgcc tgtgaagcca gctctgagga 2580
tgaggatgat gcaaccagta ccagcaatgc agaccagctg tctgaccatg gggaccttct 2640
gtctgaagag gagctggatg aatgagactc tgggaatcat ctacacagga ccaaacccaa 2700
caggcgccct ggcaccgggg aggggtagt tgtactctgc ttgtacagtc cttgagccca 2760
gtttacagat ctggagagca ggaggccagg acaaggacaa aggttgagg atggagtagg 2820
accaggggc tctgccatcc taggcacatc tcaaggctt ttatgaagac ttacagatg 2880
tcctctgtaa atagcatcga gagtggagt cagctccttt ctctactttt ttttggctg 2940
atggcacata ttattgttc tgggtctaa tcacagtgtt tctaaatgta aaaagtgc 3000
atgttgggtg agctagtccc gcgacattga gctcctctgc atgaagacac tgggtcctg 3060
catccagctg tttttattgc aaactagctc ctttctccca cactgggaac tttagtcac 3120
gaggctgtca ccacctggg agcactgggc caggctttgt agctcctgca gcagctctgc 3180
tacgtccatg tgcctcactc cagcatcat gaagctggcc cagcgccgca agtcgagttt 3240
ggtgaggtct ctggccaagg ctccagggt ctggtgcagg gacaaaggag aacacagtgc 3300
ccaaacact gggatgctct cactgctgt ggaggagag gaaacagaga cctgtagatg 3360
gatgattatt ctgccctggg actcgccaaa ctgataagga agtccaacct tagtagactt 3420
gattgtaaac tcaacaaatt tgggtattg tccccctagt acaccagtac tccagaggaa 3480
gaatgctttt cttgggagcc atagggtgaa taaaggaatg ttaactgtg aaaaaaaaaa 3540
aaaaaa

```

<210> 16  
 <211> 677  
 <212> PRT  
 <213> Homo sapiens

<400> 16  
 Met Asp Pro Glu Val Arg Leu Ser Ser Pro Pro Gly Lys Pro Glu Asp  
 1 5 10 15  
 Ser Ser Ser Val Asp Gly Gln Ser Val Gly Thr Pro Val Gly Pro Glu  
 20 25 30  
 Thr Gly Gly Glu Lys Asn Gly Pro Glu Glu Glu Glu Glu Glu Asp Phe  
 35 40 45  
 Asp Asp Leu Thr Gln Asp Glu Glu Asp Glu Met Ser Ser Ala Ser Glu  
 50 55 60  
 Glu Ser Val Leu Ser Val Pro Glu Leu Gln Glu Thr Met Glu Lys Leu  
 65 70 75 80  
 Thr Trp Leu Ala Ser Glu Arg Arg Met Ser Gln Glu Gly Glu Ser Glu  
 85 90 95  
 Glu Glu Asn Ser Gln Glu Glu Asn Ser Glu Pro Glu Glu Glu Glu Glu  
 100 105 110  
 Glu Glu Ala Glu Gly Met Glu Ser Leu Gln Lys Glu Asp Glu Met Thr  
 115 120 125  
 Asp Glu Ala Val Gly Asp Ser Ala Glu Lys Pro Pro Thr Phe Ala Ser  
 130 135 140  
 Pro Glu Thr Ala Pro Glu Val Glu Thr Ser Arg Thr Pro Pro Gly Glu  
 145 150 155 160  
 Ser Ile Lys Ala Ala Gly Lys Gly Arg Asn Asn His Arg Ala Arg Asn  
 165 170 175  
 Lys Arg Gly Ser Arg Ala Arg Ala Ser Lys Asp Thr Ser Lys Leu Leu  
 180 185 190  
 Leu Leu Tyr Asp Glu Asp Ile Leu Glu Arg Asp Pro Leu Arg Glu Gln  
 195 200 205  
 Lys Asp Leu Ala Phe Ala Gln Ala Tyr Leu Thr Arg Val Arg Glu Ala  
 210 215 220  
 Leu Gln His Ile Pro Gly Lys Tyr Glu Asp Phe Leu Gln Val Ile Tyr  
 225 230 235 240  
 Glu Phe Glu Ser Ser Thr Gln Arg Arg Thr Ala Val Asp Leu Tyr Lys  
 245 250 255  
 Ser Leu Gln Ile Leu Leu Gln Asp Trp Pro Gln Leu Leu Lys Asp Phe  
 260 265 270  
 Ala Ala Phe Leu Leu Pro Glu Gln Ala Leu Ala Cys Gly Leu Phe Glu  
 275 280 285

Glu Gln Gln Ala Phe Glu Lys Ser Arg Lys Phe Leu Arg Gln Leu Glu  
 290 295 300  
 Ile Cys Phe Ala Glu Asn Pro Ser His His Gln Lys Ile Ile Lys Val  
 305 310 315 320  
 Leu Gln Gly Cys Ala Asp Cys Leu Pro Gln Glu Ile Thr Glu Leu Lys  
 325 330 335  
 Thr Gln Met Trp Gln Leu Leu Lys Gly His Asp His Leu Gln Asp Glu  
 340 345 350  
 Phe Ser Ile Phe Phe Asp His Leu Arg Pro Ala Ala Ser Arg Met Gly  
 355 360 365  
 Asp Phe Glu Glu Ile Asn Trp Thr Glu Glu Lys Glu Tyr Glu Phe Asp  
 370 375 380  
 Gly Phe Glu Glu Val Ala Leu Pro Asp Val Glu Glu Glu Glu Glu Pro  
 385 390 395 400  
 Pro Lys Ile Pro Thr Ala Ser Lys Asn Lys Arg Lys Lys Glu Ile Gly  
 405 410 415  
 Val Gln Asn His Asp Lys Glu Thr Glu Trp Pro Asp Gly Ala Lys Asp  
 420 425 430  
 Cys Ala Cys Ser Cys His Glu Gly Gly Pro Asp Ser Lys Leu Lys Lys  
 435 440 445  
 Ser Lys Arg Arg Ser Cys Ser His Cys Ser Ser Lys Val Cys Asp Ser  
 450 455 460  
 Lys Ser Tyr Lys Ser Lys Glu Pro His Glu Leu Val Gly Ser Ser Pro  
 465 470 475 480  
 His Arg Glu Ala Ser Pro Met Pro Gly Ala Lys Glu Ala Gly Gln Gly  
 485 490 495  
 Lys Asp Met Met Glu Glu Glu Ala Pro Glu Glu Arg Glu Ser Thr Glu  
 500 505 510  
 Ala Thr Gln Ser Arg Thr Val Arg Thr Thr Arg Lys Gly Glu Met Pro  
 515 520 525  
 Val Ser Gly Leu Ala Val Gly Ser Thr Leu Pro Ser Pro Arg Glu Val  
 530 535 540  
 Thr Val Thr Glu Arg Leu Leu Leu Asp Gly Pro Pro Pro His Ser Pro  
 545 550 555 560  
 Glu Thr Pro Gln Phe Pro Pro Thr Thr Gly Ala Val Leu Tyr Thr Val  
 565 570 575  
 Lys Arg Asn Gln Val Gly Pro Glu Val Arg Ser Cys Pro Lys Ala Ser  
 580 585 590  
 Pro Arg Leu Gln Lys Glu Arg Glu Gly Gln Lys Ala Val Ser Glu Ser  
 595 600 605

Glu Ala Leu Met Leu Val Trp Asp Ala Ser Glu Thr Glu Lys Leu Pro  
 610 615 620  
 Gly Thr Val Gly Thr Pro Cys Phe Leu Pro Glu Ser Cys Phe Leu Lys  
 625 630 635 640  
 Asp Gln Arg Cys Arg Glu Lys Thr Cys Val Arg Glu Thr Arg His Ser  
 645 650 655  
 Arg Glu Met Ala Ala Leu Lys Gln Ser Ser Gly Glu Asp Lys Arg Gln  
 660 665 670  
 Asp Val Pro Cys Pro  
 675

<210> 17  
 <211> 3052  
 <212> DNA  
 <213> Homo sapiens

<400> 17  
 atgatgcgtc cctgcctcgg ccgctgcagt cgccgccgcc gccgccgcag gccgggagga 60  
 gccgcagcgc cgggcgaccc cgcccgggcc tcggatccga tcacatagga cagtatgcac 120  
 cttaaagatcc tgaagaaacg gcacaaaatg ttcaagtgat gtttagaaat aacttgtgag 180  
 ggtgcgtcag ggaatcatg cagccatcag gacacaggct ccgggacgtc gagcatcacc 240  
 ctctcctggc tgaatgatg aactatgact ctctcctgcc cctcctcctc gaggtgacg 300  
 tggctgaccg ggtctggttc atccgtgacg gctgcggcat gatctgtgct gtcattgacgt 360  
 ggcttctggg cgcttatgca gacttcgtgg tgactttcgt catgctgctg ccttccaaag 420  
 acttctggta ctctgtggtc aacggggtca tctttaactg cttggccgtg cttgcctctg 480  
 catccacact gagaaccatg ctccaccgacc ctggggcagt acccaaagga aacgctacga 540  
 aagaatacat ggagagcttg cagctgaagc ccggggaagt catctacaag tgccccaagt 600  
 gctgctgtat taaccgccag cgcgccacc actgcagtat ttgcaaaaga tgtattccga 660  
 aaatggatca tcactgccg tgggtgaaca attgtgtagg agaaaagaat caaagatttt 720  
 ttgtgctctt cactatgtat atagctctgt ctccagtcga tgctctgac ctttgtggat 780  
 ttcagttcat ctctgtgtgc cgaggcagc ggactgaatg cagtgtattt tcacctccga 840  
 taactgtaat cctgttgatc ttctgtgtgc ttgagggtct tctgttttct actttcactg 900  
 cagttatggt tggcacccaa atccactcca tatgcaacga cgagacggag atcgagcgat 960  
 tgaaaagtga gaagccaca tgggagcggg ggctgcgatg ggaagggatg aagtcctgtc 1020  
 ttggggggcc cccctcactc ctctggatga atccctttgt gggcttccga tttaggcgac 1080  
 tgccacagag acccagaaaa ggtggcccg agttctcagt gtgaggcgtg gctcatcaga 1140  
 ctgaaacttg ctacagact tccagttatt tatttgggt ctgaaggata tcaacagctc 1200  
 atctgtgacc aacagggcaa ctggaacctc cacaaccaa ttgcttgac caagcagagt 1260  
 tttatatatt tatagtcaca gatggcagag gaagaggctc tcagtcacca cctgtacaac 1320  
 aacggaaaag tgtgtggcca caggaagaag ccaaaccgct tggcctcctg cagagctggg 1380  
 gcttctgtgg agaatacttc ggggtattac atgggttatt caaatcctgg gtcctgagct 1440  
 gctgtttcca atcatgaaga aaacagtga atccagtga cagggtattc ccaagcagtc 1500  
 atttcagggg gctcctgctg acccgccac tcagcagtg actccccgga tcacagcagg 1560  
 gcgtttacat agaaagacgt tttggtctcg attagctccg atgctttgca ctgaagttgc 1620  
 aaaagatctg tgcactgaac agtgaagggt gcttcgggca cactccccgc tgccccgga 1680  
 gagacatcct ttgacctctc cagcaagtct gtgtgtgtgc gtgtctgtgc gtgtgcgcgc 1740  
 gtgtgtgcat gtgtgtcaaa attgccagt ttgtttaggc aatgtaacat ttaccggctg 1800  
 tgtacagcaa acaagctatt ttttagaac cgacgtttca gggagaggag gagagagccg 1860  
 cggggctcctg ccggtggtta ctatgaatgt attgctgttg gaggacatct cgatccaaag 1920  
 aacagccgtt cctgtgcggc ccttcgttgc cctcctgctt tcatttttta aagaaatctt 1980  
 gagtgcctga gggccttgga actgattttt tttttttgtt ccagccaaat tagcagtgtg 2040  
 taaatggcac ctaggtaaga gcagagctgc ggctcggtga cttgatactt gggcagccc 2100  
 gatgctctgt gtggggcagg ggagggcatcc ttactggaga ggcagggccc agccattggg 2160  
 cacctctggg aaggggaggg gaccatgagg cagccagccc ctggcagggg cgactgtgcc 2220



```

accgcaggca gcgctccagt cggaatggc caggatggcg ccctcttgtt ggagtttttg 2280
gttagctttt acgttttttt ctccaccac ggacacagtg ataaaatagg atccttggtg 2340
cggagcttaa aattatgcca gaaagccaac agctcccttc gtggggcctt gccttaaact 2400
tgcttggttt gtacattttt tgccggacgc atcaagaagc aatctgtgac aaagtctgag 2460
ggtcttcctt tatgcttgcc ctccacatta agagaagttg ggtctccctc ctgggaattg 2520
tttggccttt ctgttcattt gtgaactggt ttttgttttt aattactctg taccatcc 2580
gaatcagggc ttctaccact gctgatgcaa aaccacaaag ggacctacct gagccaccgt 2640
cctagccaag cgagcaaacc tgcagggggg ttggaagtgg acttggtcac cgcagaagcg 2700
tgtgcgccgt tgggggaaga gctgcgtcac agccagaggg acaaagtgtg ggtgatcctg 2760
gagacgccag ttctcgagat tgctctgcat attcatttgc acattgttgt ctgggttgga 2820
catgcgtgtg ggtcttcagt tgaggctttt aatatgtata tcctgttatc aataaaacaa 2880
ttatccaagt ggttgaatcc tgtgagactt ggcaagtgtg tgcaaatcaa gtataactga 2940
cttttcaacc tcttctttca atgtaacttt tatatgaaat aaagtaatca attaacagtt 3000
ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 3052

```

&lt;210&gt; 18

&lt;211&gt; 308

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 18

```

Met Gln Pro Ser Gly His Arg Leu Arg Asp Val Glu His His Pro Leu
  1              5              10              15

Leu Ala Glu Asn Asp Asn Tyr Asp Ser Ser Ser Ser Ser Ser Ser Glu
          20              25              30

Ala Asp Val Ala Asp Arg Val Trp Phe Ile Arg Asp Gly Cys Gly Met
  35              40              45

Ile Cys Ala Val Met Thr Trp Leu Leu Val Ala Tyr Ala Asp Phe Val
  50              55              60

Val Thr Phe Val Met Leu Leu Pro Ser Lys Asp Phe Trp Tyr Ser Val
  65              70              75              80

Val Asn Gly Val Ile Phe Asn Cys Leu Ala Val Leu Ala Leu Ser Ser
          85              90              95

His Leu Arg Thr Met Leu Thr Asp Pro Gly Ala Val Pro Lys Gly Asn
 100              105              110

Ala Thr Lys Glu Tyr Met Glu Ser Leu Gln Leu Lys Pro Gly Glu Val
 115              120              125

Ile Tyr Lys Cys Pro Lys Cys Cys Cys Ile Lys Pro Glu Arg Ala His
 130              135              140

His Cys Ser Ile Cys Lys Arg Cys Ile Arg Lys Met Asp His His Cys
 145              150              155              160

Pro Trp Val Asn Asn Cys Val Gly Glu Lys Asn Gln Arg Phe Phe Val
 165              170              175

Leu Phe Thr Met Tyr Ile Ala Leu Ser Ser Val His Ala Leu Ile Leu
 180              185              190

Cys Gly Phe Gln Phe Ile Ser Cys Val Arg Gly Gln Trp Thr Glu Cys
 195              200              205

```

Ser Asp Phe Ser Pro Pro Ile Thr Val Ile Leu Leu Ile Phe Leu Cys  
 210 215 220

Leu Glu Gly Leu Leu Phe Phe Thr Phe Thr Ala Val Met Phe Gly Thr  
 225 230 235 240

Gln Ile His Ser Ile Cys Asn Asp Glu Thr Glu Ile Glu Arg Leu Lys  
 245 250 255

Ser Glu Lys Pro Thr Trp Glu Arg Arg Leu Arg Trp Glu Gly Met Lys  
 260 265 270

Ser Val Phe Gly Gly Pro Pro Ser Leu Leu Trp Met Asn Pro Phe Val  
 275 280 285

Gly Phe Arg Phe Arg Arg Leu Pro Thr Arg Pro Arg Lys Gly Gly Pro  
 290 295 300

Glu Phe Ser Val  
 305

<210> 19  
 <211> 3123  
 <212> DNA  
 <213> Homo sapiens

<400> 19  
 ggatgctgga gcagaacaat ggattttctct ttctctttca tgcaagggat catgggaaac 60  
 acaattcagc aaccacctca actcattgac tccgccaaaca tccgtcagga ggatgccttt 120  
 gataacaaca gtgacattgc tgaagatggg ggccagacac catatgaagc tactttgcag 180  
 caaggctttc agtaccagc tacaacagaa gatcttcctc cactcacaaa tgggtatcca 240  
 tcatcaatca gtgtgtatga aactcaaac aaataccagt catataatca gtatccta 300  
 gggtcagcca atggcttttg tgcagttaga aacttttagc cactgacta ttatcattca 360  
 gaaattccaa acacaagacc acatgaaatt ctggaaaaac cttccctcc acagccacca 420  
 cctcctcctt cgggtaccaca aactgtgatt ccaaagaaga ctggctcacc tgaaattaaa 480  
 ctaaaaataa ccaaaactat ccagaatggc aggaattgt ttgagtctt cctttgtgga 540  
 gaccttttaa atgaagtaca ggcaagtgg caccgaaat caaagcatga aagcagaaaa 600  
 gaaaagagga aaaaagcaa caagcatgac tcatcaagat ctgaagagcg caagtcacac 660  
 aaaaatccca aattagaacc agaggaacaa aatagaccaa atgagagggg tgacactgta 720  
 tcagaaaaac caaggaaga accagtacta aaagaggaag cccagttca gccataacta 780  
 tcttctgttc caacaacgga agtggtccact ggtgttaagt ttcaggttg cgatcttg 840  
 tgggtccaagg tgggaaccta tccttggtgg ccttgatgg tttcaagtga tcccagctt 900  
 gaggttcata ctaaaattaa cacaagagg gcccagaaat atcatgtcca gtttttttagc 960  
 aaccagccag agagggcggt ggttcatgaa aaacgggtac gagagtataa aggtcataaa 1020  
 cagtatgaag aattactggc tgaggcaacc aaacaagcca gcaatcactc tgagaacaa 1080  
 aagattcggg aaccccgacc tcagagagaa cgtgctcagt gggatattgg cattgccc 1140  
 gcagagaaag cattgaaat gactcgagaa gaaagaatag aacagtatac ttttatttac 1200  
 attgataaac agcctgaaga ggctttatcc caagcaaaaa agagtgttgc ctccaaaacc 1260  
 gaagttaaaa aaacccgacg accaagatct gtgctgaata ctcagccaga acagaccaat 1320  
 gcaggggagg tggcctctc actctcaagt actgaaatc ggagacatag ccagaggcgg 1380  
 cacacaagtg cggaagagga agagccaccg cctgttaaaa tagcctggaa aactgcggca 1440  
 gcaaggaat ccttaccagc ttccattacg atgcacaaag ggagcctgga ttgacagaag 1500  
 tgaacatgt ctccagttgt gaaaattgaa caagtgtttg ctcttcagaa tgctacagg 1560  
 gatgggaaat ttatcgatca atttgtttat tcaacaaagg gaattggtaa caaacagaa 1620  
 ataagtgtca gggggcaaga caggcttata atttctacac caaacagag aaatgaaaag 1680  
 ccaacgcaga gtgtatcatc tcctgaagca acatctggtt ctacaggctc agtagaaaag 1740  
 aagcaacaga gaagatcaat tagaactcgt tctgaatcag agaaatccac tgaggttgtg 1800  
 ccaaagaaga agatcaaaaa ggagcaggtt gaaacagttc ctcaggctac agtgaagact 1860  
 ggattacaga aagggtcggc ggaccgggga gtgcagggct ctgtcagatt cagtgcagc 1920

```

tccgtctccg cagcgattga ggaaactgtg gactgagatt cctgtacaat ttcaccccag 1980
aaactccaga cttgtagtct ccatgcaaga tttctttgtc ggcggcttga taaacagttt 2040
ctttgttttc gattttgatt tcgccaatca tcattattgg cattttcctg cctggtttct 2100
tcttcaagac tctgaacaat tgctttaaca gtcaaatgat ttttttttcc ggttttgagct 2160
ggatgggtac agcttaaatc atgggtccag cctaaaaacc accatttaac ttacactgat 2220
caatttcaac atggactgtt tttggttttt tgttttttaa taaagcatca ttaatgcaca 2280
tctgcagggg tttgccaaac agcccaaact gtatacatta caatcattaa aagtctctat 2340
tttttttaat attagtgcg ttatcatgga gaacagcatg acagctgtct ttggcagtct 2400
gtcatttttc tagcattttc agaaactcat cggaaatggc ggtacctgtg tttcccttcg 2460
aaagcctctc agtacagcac tcctgttcct ctgttaaaac tccttgtaa tccagtgate 2520
ttttaggcca aggaaatatt ttgtgatggt gttctgggtc catacaccag caatgaagga 2580
gatagatttg tgtacttggt ttttttaatc agcattaaca tgggcaggca ccctcattta 2640
tagatgtcag gaaacattca gtgaaaaact tgtagaatgg gatgtgataa cgagggtcca 2700
gtaatctgag cagtctaacg agggccacct cctccaccac agaactgtggc tatgttccaa 2760
gtgctactct cactcagcct gttgcggatc ttcatggcct caggagactt gttctccat 2820
gggctcttct ggactgcaca cttccaccat agcttgctgg gttgatctag atgtctgttt 2880
gttgatgga aattttgggg gaaaaaatcc aaaacacaaa ctgtgggttg aaatattaac 2940
cgtctccttg gttccttggt attcaccgtg cctgatctgc acatttcac gtggctgttt 3000
ctgtatagcc tatactgcat tagcccaaga gattgttgc ttgtaacttt ttgcactatt 3060
gttttggtg gatttgatt acacacagtt ttaaaaaaaaa caataaaaaa aaaaaaaaaa 3120
aaa
3123

```

<210> 20  
 <211> 645  
 <212> PRT  
 <213> Homo sapiens

```

<400> 20
Met Asp Phe Ser Phe Ser Phe Met Gln Gly Ile Met Gly Asn Thr Ile
  1             5             10            15

Gln Gln Pro Pro Gln Leu Ile Asp Ser Ala Asn Ile Arg Gln Glu Asp
  20            25            30

Ala Phe Asp Asn Asn Ser Asp Ile Ala Glu Asp Gly Gly Gln Thr Pro
  35            40            45

Tyr Glu Ala Thr Leu Gln Gln Gly Phe Gln Tyr Pro Ala Thr Thr Glu
  50            55            60

Asp Leu Pro Pro Leu Thr Asn Gly Tyr Pro Ser Ser Ile Ser Val Tyr
  65            70            75            80

Glu Thr Gln Thr Lys Tyr Gln Ser Tyr Asn Gln Tyr Pro Asn Gly Ser
  85            90            95

Ala Asn Gly Phe Gly Ala Val Arg Asn Phe Ser Pro Thr Asp Tyr Tyr
 100            105            110

His Ser Glu Ile Pro Asn Thr Arg Pro His Glu Ile Leu Glu Lys Pro
 115            120            125

Ser Pro Pro Gln Pro Pro Pro Pro Pro Ser Val Pro Gln Thr Val Ile
 130            135            140

Pro Lys Lys Thr Gly Ser Pro Glu Ile Lys Leu Lys Ile Thr Lys Thr
 145            150            155            160

Ile Gln Asn Gly Arg Glu Leu Phe Glu Ser Ser Leu Cys Gly Asp Leu
 165            170            175

```

Leu Asn Glu Val Gln Ala Ser Glu His Thr Lys Ser Lys His Glu Ser  
 180 185 190  
 Arg Lys Glu Lys Arg Lys Lys Ser Asn Lys His Asp Ser Ser Arg Ser  
 195 200 205  
 Glu Glu Arg Lys Ser His Lys Ile Pro Lys Leu Glu Pro Glu Glu Gln  
 210 215 220  
 Asn Arg Pro Asn Glu Arg Val Asp Thr Val Ser Glu Lys Pro Arg Glu  
 225 230 235 240  
 Glu Pro Val Leu Lys Glu Glu Ala Pro Val Gln Pro Ile Leu Ser Ser  
 245 250 255  
 Val Pro Thr Thr Glu Val Ser Thr Gly Val Lys Phe Gln Val Gly Asp  
 260 265 270  
 Leu Val Trp Ser Lys Val Gly Thr Tyr Pro Trp Trp Pro Cys Met Val  
 275 280 285  
 Ser Ser Asp Pro Gln Leu Glu Val His Thr Lys Ile Asn Thr Arg Gly  
 290 295 300  
 Ala Arg Glu Tyr His Val Gln Phe Phe Ser Asn Gln Pro Glu Arg Ala  
 305 310 315 320  
 Trp Val His Glu Lys Arg Val Arg Glu Tyr Lys Gly His Lys Gln Tyr  
 325 330 335  
 Glu Glu Leu Leu Ala Glu Ala Thr Lys Gln Ala Ser Asn His Ser Glu  
 340 345 350  
 Lys Gln Lys Ile Arg Lys Pro Arg Pro Gln Arg Glu Arg Ala Gln Trp  
 355 360 365  
 Asp Ile Gly Ile Ala His Ala Glu Lys Ala Leu Lys Met Thr Arg Glu  
 370 375 380  
 Glu Arg Ile Glu Gln Tyr Thr Phe Ile Tyr Ile Asp Lys Gln Pro Glu  
 385 390 395 400  
 Glu Ala Leu Ser Gln Ala Lys Lys Ser Val Ala Ser Lys Thr Glu Val  
 405 410 415  
 Lys Lys Thr Arg Arg Pro Arg Ser Val Leu Asn Thr Gln Pro Glu Gln  
 420 425 430  
 Thr Asn Ala Gly Glu Val Ala Ser Ser Leu Ser Ser Thr Glu Ile Arg  
 435 440 445  
 Arg His Ser Gln Arg Arg His Thr Ser Ala Glu Glu Glu Glu Pro Pro  
 450 455 460  
 Pro Val Lys Ile Ala Trp Lys Thr Ala Ala Ala Arg Lys Ser Leu Pro  
 465 470 475 480  
 Ala Ser Ile Thr Met His Lys Gly Ser Leu Asp Leu Gln Lys Cys Asn  
 485 490 495

Met Ser Pro Val Val Lys Ile Glu Gln Val Phe Ala Leu Gln Asn Ala  
 500 505 510  
 Thr Gly Asp Gly Lys Phe Ile Asp Gln Phe Val Tyr Ser Thr Lys Gly  
 515 520 525  
 Ile Gly Asn Lys Thr Glu Ile Ser Val Arg Gly Gln Asp Arg Leu Ile  
 530 535 540  
 Ile Ser Thr Pro Asn Gln Arg Asn Glu Lys Pro Thr Gln Ser Val Ser  
 545 550 555 560  
 Ser Pro Glu Ala Thr Ser Gly Ser Thr Gly Ser Val Glu Lys Lys Gln  
 565 570 575  
 Gln Arg Arg Ser Ile Arg Thr Arg Ser Glu Ser Glu Lys Ser Thr Glu  
 580 585 590  
 Val Val Pro Lys Lys Lys Ile Lys Lys Glu Gln Val Glu Thr Val Pro  
 595 600 605  
 Gln Ala Thr Val Lys Thr Gly Leu Gln Lys Gly Ser Ala Asp Arg Gly  
 610 615 620  
 Val Gln Gly Ser Val Arg Phe Ser Asp Ser Ser Val Ser Ala Ala Ile  
 625 630 635 640  
 Glu Glu Thr Val Asp  
 645

<210> 21  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> oligonucleotide

<220>  
 <221> misc\_feature  
 <222> (2)  
 <223> biotinylated phosphoramidite residue

<400> 21  
 gntggcattc caggtgtggt ggtgcacag

29

<210> 22  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> oligonucleotide

<220>  
 <221> misc\_feature  
 <222> (2)  
 <223> biotinylated phosphoramidite residue

<400> 22  
cnaggcta at gaggcagcaa tggcttaaa 29

<210> 23  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<400> 23  
tcttcactgt atctcctttg c 21

<210> 24  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>  
<221> misc\_feature  
<222> (2)  
<223> biotinylated phosphoramidite residue

<400> 24  
anaactccca gaaggaagga gatggatcc 29

<210> 25  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>  
<221> misc\_feature  
<222> (2)  
<223> biotinylated phosphoramidite residue

<400> 25  
anattcttgt gtagaagcga actgtccag 29

<210> 26  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>  
<221> misc\_feature  
<222> (2)  
<223> biotinylated phosphoramidite residue

<400> 26  
anggcgaaga agatgagcgc ggcagtgag 29

<210> 27  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>  
<221> misc\_feature  
<222> (2)  
<223> biotinylated phosphoramidite residue

<400> 27  
gnaagaagtt taagggaata aggagtgt 29

<210> 28  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>  
<221> misc\_feature  
<222> (2)  
<223> biotinylated phosphoramidite residue

<400> 28  
cncctttttcc tcttggttctt tgaggctgt 29

<210> 29  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>  
<221> misc\_feature  
<222> (2)  
<223> biotinylated phosphoramidite residue

<400> 29  
anggatcaga gcatggactg aagacagag 29

<210> 30  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> oligonucleotide

<220>

<221> misc\_feature  
 <222> (2)  
 <223> biotinylated phosphoramidite residue

<400> 30  
 anatagacca aatgagaggg ttgacactg

29

<210> 31  
 <211> 87  
 <212> PRT  
 <213> Homo sapiens

<400> 31  
 Met Glu Met Gly Ile Gln Val Gly Lys Arg Leu Leu Ile Gln Gly Leu  
   1                  5                  10                  15  
 Glu Leu Gly Glu Leu Leu Thr Lys Cys Gly Asn Glu Asp Phe Leu Lys  
                   20                  25                  30  
 Met Gly Ala Gly Trp Ile Gln Ala Ser Tyr Phe Asn Pro Gly Leu Thr  
           35                  40                  45  
 Leu Ile Gln Ser Lys Tyr Phe Leu Ser Ser Ser Trp Gly Gly Val Leu  
   50                  55                  60  
 Gly Gln Glu Lys Lys Gly Asp Lys Asn Thr Ile Pro Glu Lys Glu Glu  
   65                  70                  75                  80  
 Lys Tyr Leu His Phe Val Leu  
                   85

<210> 32  
 <211> 82  
 <212> PRT  
 <213> Homo sapiens

<400> 32  
 Met Met Phe Ile Val Pro Val Val Phe His Pro Arg Ser Arg Ile Ser  
   1                  5                  10                  15  
 Phe Ser Thr Phe Ser Phe Ile Arg Val Met Lys Leu Asn Pro Trp Ala  
           20                  25                  30  
 Met Ser Glu Ala Gln Ser Leu Glu Cys Val Tyr Ser Gln Trp Cys Met  
           35                  40                  45  
 Tyr Ile Leu Cys Leu Asp Ser Leu Arg Ser Val Ser Glu Asn Leu Asp  
   50                  55                  60  
 Ser Ser Leu Leu His Lys Asn Phe Ile Cys Ile Tyr Glu Asp Asp Ser  
   65                  70                  75                  80  
 Val Pro



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/06946

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 15/00, 15/12; CO7K 14/00, 14/435, 7/00

US CL : 435/69.1, 320.1, 325; 536/23.5, 23.1; 530/350, 300

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 320.1, 325; 536/23.5, 23.1; 530/350, 300

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Genbank, GenEmbl, geneseq, Swissprot, pir58

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TOUCHMAN et al. "2006 Expressed-Sequence Tags derived from human chromosome 7-enriched cDNA libraries." Genome Research. 1997, Vol.7, No.3, pages 281-292, especially Figure 3 and Genbank Accession No. G30725.	1-3
X	Database on Genbank, HILLIER et al. "ym51g12.s1 Homo sapiens cDNA clone 51970 3'." 06 July 1995, Accession No. H23117 and 23116	1-3
X	Database on Genbank, HILLIER et al. "yg2605.r1 Homo sapiens cDNA clone 33723 5'." 14 April 1995, Accession No. R19556.	
X	Database on Genbank, HILLIER et al. "yj85c08.r1 Homo sapiens cDNA clone 155534 5'." 02 June 1995, Accession No. R71701.	1-3

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

15 JUNE 1999

Date of mailing of the international search report

16 JUL 1999

 Name and mailing address of the ISA/US  
 Commissioner of Patents and Trademarks  
 Box PCT  
 Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

  
 ELIANE LAZAR-WESLEY

Telephone No. (703) 308-0196

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US99/06946

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-11

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.